Towards 'Ubiquitous' Ubiquitous Computing: an alliance with 'the Grid'

Oliver Storz, Adrian Friday, and Nigel Davies

Computing Department,
Lancaster University,
Lancaster, UK
{oliver,adrian,nigel}@comp.lancs.ac.uk

Abstract. The drive to implement Ubiquitous Computing systems has driven the development of numerous research prototypes. Distributed systems platforms have emerged, each addressing a subset of the overall problem space. In contrast, many other scientific disciplines have successfully embraced a standardised vision of a global distributed computing platform, i.e. the Grid. While the Grid certainly is far from perfect platform, in this position paper we argue that significant potential exists for building ubiquitous computing applications on a hitherto unprecedented scale. Additionally, we explore the areas of synergy between Grid computing and ubiquitous computing and highlight a number of common research challenges, and argue why this platform should be used to accelerate the deployment of ubiquitous computing.

1 Introduction

When Weiser inspired researchers all over the world by espousing his vision of ubiquitous computing (Ubicomp) [1], he expected a gradual movement towards embodied virtuality within 20 years, predicting that this would emerge as the dominant form of human computer interaction. We observe that, despite numerous lab-based prototypes, ubiquitous computing has thus far failed to become commonplace.

Many reasons have been cited for this slow progress [2, 3], including the lack of common distributed systems infrastructure. While many researchers have begun to target this space: e.g. GAIA [4], Aura [5] and Cooltown [6], such platforms have failed to become widely adopted as a common standard. This may be in part due to the difficulty in persuading others to use proprietary technologies whose future development and support is unclear, yet, even where these platforms have attempted to leverage existing standards (e.g. Cooltown's extensive use of Web protocols) they still have not experienced significant uptake.

In contrast, within the context of other scientific disciplines we are witnessing the evolution of a vision of global distributed computing: *the Grid* [7]. The Grid promises a world where access to (computational) resources across institutional boundaries is standardised, uniform, inexpensive, ubiquitous and reliable (see section 2). While the origins of this vision might not exactly match that of ubiquitous computing, many common elements can be identified. In this paper we argue that adoption of a common standard, such as the Grid, is essential for realising ubiquitous computing on a wider scale. Moreover, we speculate that through active participation in Grid standardisation, the Grid platform may even become the standard for engineering ubiquitous computing environments.

2 The Grid Vision

The Grid is often seen as a platform for networking large computational resources to support execution of demanding scientific experiments. Indeed, the aim of the Grid is to make processing power and data storage ubiquitous, analogous to the electricity distribution grid. Ian Foster, who is given most of the credit for coining the term *Grid*, describes a computational Grid in [7] as

> "a hardware and software infrastructure that provides dependable, consistent, pervasive and inexpensive access to high-end computational capabilities."

Over time, however, the Grid has moved away from the mere sharing of *computational* resources and has become a more generic platform for the sharing of *any kind* of networked resource. Foster therefore recently described the the Grid as a system that

> "coordinates resources that are not subject to centralised control ... using standard, open, general-purpose protocols and interfaces ... to deliver non-trivial qualities of service." [8]

If we compare Foster's vision with that of Weiser's, we begin to find similarities: in ubiquitous computing we speak of augmenting the environment with large numbers of devices and services; these devices will *not* be *under centralised control*. Moreover, in order to create valuable services to the human user, these services and devices will have to interact in a more or less *coordinated* fashion and these devices will be heterogeneous in nature, making it necessary to specify *standardised* ways for interaction. If we redefine the notion of 'a resource' to include services and devices, we can clearly see that the basic principles of ubiquitous computing and Grid computing are not very far apart. Indeed, the Grid is beginning to embrace a wider range of emerging scientific disciplines (such as pervasive sensing and medical informatics [9]).

Inter-organisational sharing of resources is not a task that can be achieved easily and many hard computer science problems remain to be addressed [10, 7]. Comparing these issues against the basic requirements of ubiquitous computing (as recently articulated by [2, 3]), reveals a substantial synergy, as we enumerate below:

– **Heterogeneity and Interoperability**: Both ubiquitous computing and the Grid involve a large number of heterogeneous resources. Standardised mechanisms for inter-resource communication have to be defined. These

mechanisms have to be extensible for the seamless inclusion of future resources.
- **Scalability**: Grid applications typically consist of highly parallel or massively replicated computations. In ubiquitous computing, the environments themselves are comprised of innumerable devices, sensors and computational services. Both environments offer similar challenges of scalability, as the number of resources/services increases, adapts and evolves to encompass new nodes and users.
- **Adaptability and Fault Tolerance**: The Grid and ubiquitous computing environments are both highly complex distributed environments. Both kinds of environment will be required to cope with change, failure and the introduction of new components – gracefully and at run-time – it is not practical to reboot such systems to fix faults or return to a particular known-good state! While the semantics of failure are not yet well defined in ubiquitous computing, it is clear that platforms must offer the ability to adapt to changes in their underlying environment and offer dependable failure modes.
- **Resource Management and Service Composition**: Clearly related to scalability is resource management and higher level service composition: as Grid services become commonplace and the range of services on offer diversifies, meta-level services and tools will be required to help automate and control the life-cycle, interaction, monitoring and dynamic composition of simple services to form higher level applications [11] and produce content. Witness the difficulties in emerging ubiquitous computing environments in writing portable "applications" which run *across* more than one environment.
- **Service Discovery**: In both application domains, effective and efficient ways of discovering services and resources need to be developed. This is also an essential requirement for the efficient deployment of resource management and service composition mechanisms.
- **Security**: The Grid and ubiquitous computing both need consistent security architecture that is scalable and lightweight. There are many non-trivial issues to address in this arena: large numbers of users must be *authenticated* and *authorised* to access resources, without relying on a centralised infrastructure; perhaps, utilising techniques such as *recognition*, *reputation* and *trust*. Ensuring data *integrity*, *confidentiality* and *privacy* are other aspects that have only begun to be addressed: indeed, protecting users from unwanted disclosures of private or confidential aspects of their interactions within ubicomp offers some unique, and possibly intractable challenges (e.g. due to the pervasive use of embedded sensing). Furthermore, any new security techniques need to be capable of integrating into organisations' existing security and administrative structures, with due consideration to both the technical and social issues this implies [12].
- **Communication**: Services/resources may be interconnected various forms of communication , impacting on higher layer protocols. Current Grid protocols assume plentiful bandwidth and reliable, always available communications – features not typically found in wireless networks. Some applications

and services demand specific Quality of Service (e.g. timely delivery of significant events, or of streaming media). Means for specifying these requirements have to be supplied, as well as underlying mechanisms for enforcing them. Already a challenging problem in conventional networks, it is likely that ubiquitous computing will raise additional challenges due to the variety of devices and many possible paths interconnection.

- **Audit Trails**: Applications in both ubiquitous computing and the Grid will involve interactions with large numbers of heterogeneous resources. It is vital to provide means for making sense of these underlying system processes to support users (e.g. for reassurance or simply to support application development). Making technology "calm" will inevitably involve the delegation of decisions away from the user into "the environment" [13]. However, from time to time, users might still be interested in the steps and decisions that have lead to the initiation of a particular action, e.g. in the case of a "misbehaving" smart room. In the Grid context, researchers might be similarly interested in the progress of a large-scale highly distributed computation.
- **Payment**: In both computing domains, the deployment of infrastructure is an undertaking that can be very costly. Means have to be found for financing these infrastructures, for example by directly imposing charges for accessing and using services and resources or via cross-subsidisations from other sectors [14].

As we can see, the requirements for building computational Grids significantly overlap with those for supporting ubiquitous computing applications. However, currently there has been little recognition in the Grid community of the needs of ubiquitous computing. Recent documents, such as [15, 16] still focus on computational resources and their possible exploitation within a global Grid infrastructure. The logical consequence is that there is little sensitivity within the Grid research community to problems that are specific and crucial to ubiquitous computing, such as the limitations of light-weight, possibly embedded devices.

3 Using Grid Technologies

So far we have only been talking about *the vision* of the Grid. In this section we consider to what extent Grid middleware is usable for building pervasive computing applications.

Although there already exist a variety of commercial and non-commercial Grid middleware platforms, e.g. Globus v2 [17], Legion [18] and Avaki [19], one can identify a strong movement throughout both sectors towards one single architecture, the Open Grid Services Architecture (OGSA) [16]. OGSA is a joint effort to agree on a standardised next-generation platform for future Grids driven by the Global Grid Forum (GGF) [20]. With the adoption of OGSA, the global Grid community has accepted and realised one of the main visions of Grid computing: to agree upon open standards that provide the means for interoperability between resources. Although still undergoing standardisation, OGSA already provides some basic functionality for building computational Grids, including introspection, registration, eventing, lifecycle management, service creation and naming. Standard interfaces for these services and their behaviour are in the process of being defined as part of the Open Grid Services Infrastructure (OGSI) [21]. OGSI is commonly seen as a foundation for providing the infrastructure for building OGSA.

Being work in progress, OGSI and OGSA are far from complete, and arguably, still far removed from providing the infrastructure necessary for building ubiquitous computing. Nevertheless, we have begun to explore whether these platforms offer enough functionality by building small prototype ubiquitous computing applications. We outline one example of this approach in the following scenario (taken from one of a number of new projects [[9], [22]] that are putting new classes of device, sensor and information on the Grid).

> "Since his birth, John has been suffering from chronic lung disease, which requires that he undergoes lifelong oxygen treatment. To help him cope, John has been given a small wearable computer that records his vital signs; blood pressure, pulse and, most importantly, the oxygen saturation in his blood. John's wearable is equipped with wireless communication facilities that are used for periodically transmitting the recorded data to a small number of software components. One of these services is responsible for storing the data in a large distributed database. The database can be accessed by clinicians to track developments in John's condition. The database also contains information about the state of John's surroundings, e.g. temperature and the quality of the air. This type of data gets automatically collected from all environmental sensors that are in John's immediate vicinity. A second software component processes John's data as it enters the database, looking for anomalies in his vital signs that might necessitate immediate action. Once an anomaly has been detected, an expert system assesses the severity of the event. In case of an emergency, John's practitioner is automatically contacted and the system dispatches an emergency response team to John's current location. In case of minor events that can be handled by John himself, a notification is sent to his wearable computer which will then inform its wearer."

This scenario is far from being science-fiction – a proof-of-concept prototype based on a sensor jacket developed by partners specialising in medical sensing. The jacket allows unobtrusive monitoring of children suffering from chronic medical disorders. Most of the components of our demonstrator have been engineered as Grid services using an early release of the Globus Toolkit version 3 (GT3). Our implementation relies on OGSI's eventing framework. Future evolutions of the demonstrator will be extended to use OGSI's registry services for service discovery and the database access components of the OGSA Database Access and Integration framework OGSA-DAI [23]. OGSA-DAI will provide standardised access to existing heterogeneous databases. One can easily imagine making use of compute resources to process and visualise the data being gathered from large numbers of field experiments and clinical trials in real-time, e.g. for finding correlations, epidemiology and detecting trends in the data gathered from communities of patients suffering from the same diseases. One can also clearly envision how these widely deployed sensor networks and vast repositories of information can be re-purposed to support new experiments and, significantly, new ubiquitous computing applications.

4 Discussion

The Grid as an ideal. Using the Grid platform to support our work within ubiqutious computing environments is clearly contentious – we will return to this point again in a moment. Firstly, let us discuss one of the key challenges we face as a community: migrating from point-sample exploratory ubicomp prototypes to more widespread, and widely used systems. We do not argue that the world will become a homogenised place with one, uniform, all pervading ubicomp platform – the world will always be filled with political, social and practical limitations that will partition, segregate and divide, obviating any attempts in move in this direction.

Instead, we argue that ubicomp will only be achieved by harmonisation and integration. This requires ubicomp researchers to do two things; agree on common mechanisms to permit interworking (standardisation), and identify the core services that comprise ubicomp environments (standard services with well defined interfaces). The actual technologies chosen to underpin these decisions are not important, providing that a technology is chosen. However, we believe that there are a number of advantages for choosing the Grid.

- **Resources.** The Grid is populated with the output of expensive research programmes: it has been invested with computational, storage and human resources on a scale not possible within individual research programmes. By adopting the Grid we potentially gain access to these.
- **Context.** New Grid initiatives are providing widely deployed sensors and information systems. These sources will provide standardised access to valuable resources otherwise beyond the scope of most projects (e.g. weather monitoring, environmental sensing, telemetry data).
- **Community.** The Grid (and e-Science programmes based on Grid technology) offer access to other communities both within other scientific domains and as users. These can be seen as both access to additional intellectual and knowledge resources, and as access to potential users of ubiquitous computing, providing real requirements and dissemination possibilities for our research.
- **Politics.** Lastly, and in some ways most significantly, the Grid has political capital. One cannot underestimate the potential impact of the 'political will' behind the Grid – such has the potential to ensure research is suitably resourced and that new information and services will continue to appear on

the Grid (some of which will only be possible with buy in from the political infrastructure, e.g. traffic sensing systems, pervasive environmental or healthcare monitoring).

We also see the Grid as offering wider scope for deployment of ubicomp applications: pooling our knowledge and resources with others will sensitise us to issues of collaboration and scalability, and permit testing on a wider scale – in short, moving beyond isolated lab-based tests under artificial conditions and making us conscious of the issues related to re-usability and deployment.

The Grid as a technology. As to whether the Grid middleware as it currently stands is the ideal choice for helping us build and integrate the above services, we have found GT3 to be a viable, if rudimentary, platform for exploring the services needed within the context of reactive 'smart' environments [24] and medical monitoring applications [9]. GT3 has enabled the construction of a basic testbed without the effort of building our own middleware. Grid services, leveraging underlying and well established web protocols, *do* provide an appropriate mechanism for defining interfaces to services. The unparalleled acceptance of the web protocols and well understood solutions to running them across administrative domains, also make this one of the better choices.

We have, however, been forced to acknowledge that even in our limited experience, the functionality provided by GT3, OGSI and OGSA is far from sufficient to support all of our needs (a mere a subset of those in ubiquitous computing as a whole) – the existing service discovery mechanisms (registry based) do not facilitate dynamic deployment (it is also not clear how clients will discover registries themselves, although this is a simple bootstrapping problem). The platform does not cope well with partial connectivity nor easily support the integration of lightweight components (e.g. wearable or sensor platforms).

Although we believe the Grid can not be seen as the perfect ubiquitous computing platform, we are convinced that the time has come for the ubiquitous computing community to become actively involved in the process of evolving Grid standards. The Grid seems to have reached an important turning point: until now, most of the deployments of Grid middleware mainly dealt with computational tasks present in other science domains, defining a particular set of requirements. To support our community in achieving ubiquitous computing on a wider scale, we must get involved to help raise the profile of needs of ubiquitous computing within the Grid. Furthermore, with the arrival of OGSA we have seen the Grid move towards a more service-oriented environment. We believe that by focusing on the lessons learnt in our research environments and refining these to provide the core services required for ubiquitous computing (rather than focusing on our own unique middlewares), we will be able to achieve 'ubiquitous' ubiquitous computing.

5 Conclusion

While significant progress has already been made in the area of ubiquitous computing, there still is no standardised, uniform platform, preventing wide-scale deployment of applications in this area. We have argued that the Grid could provide a viable route for accelerating the deployment of ubiquitous computing, by helping us to define the core services and interfaces that comprise it, and permitting access to resources not possible in isolation.

We also argue that the Grid could form the basis of a platform for ubiquitous computing: both to link systems and resources together, and provide allies in addressing the numerous hard computing problems that await in trying to deploy ubicomp on a wide scale. The Grid and ubiquitous computing share many common requirements, and we argue that working together in this regard, will afford additional opportunities.

Lastly, we have identified that Grid research is at a turning point that may well influence the usability of future Grid platforms for the purposes of ubiquitous computing. Although current Grid toolkits do not yet allow us to build sophisticated applications, our first applications have demonstrated that it is already possible to use Grid technology for prototyping small ubiquitous computing applications. We will therefore continue to use and explore Grid technologies for building ubiquitous computing applications 'in the large'.

Acknowledgements

We would like to acknowledge the EPSRC in supporting our work in the Grid based Medical Devices for Everyday Health: GR/R85877 and Equator GR/N15986 projects.

References

1. Weiser, M.: The computer for the 21st century. Scientific American (1991) 94–104
2. Davies, N., Gellersen, H.: Beyond prototypes: Challenges in deploying ubiquitous systems. IEEE Pervasive Computing **1** (2002) 26–35
3. Kindberg, T., Fox, A.: System software for ubiquitous computing. IEEE Pervasive Computing **1** (2002) 70–81
4. Romn, M., Hess, C., Cerqueira, R., Ranganathan, A., Campbell, R., Nahrstedt, K.: Gaia: A middleware infrastructure to enable active spaces. IEEE Pervasive Computing **1** (2002)
5. Sousa, J., Garlan, D.: Aura: An architectural framework for user mobility in ubiquitous computing environments. In Bosch, J., Gentleman, M., Hofmeister, C., Kuusela, J., eds.: Software Architecture: System Design, Development, and Maintenance (Proceedings of the 3rd Working IEEE/IFIP Conference on Software Architecture), Kluwer Academic Publishers (2002) 29–43
6. Kindberg, T., Barton, J., Morgan, J., Becker, G., Caswell, D., Debaty, P., Gopal, G., Frid, M., Krishnan, V., Morris, H., Schettino, J., Serra, B., Spasojevic, M.: People, places, things: web presence for the real world. Mobile Networks and Applications **7** (2002) 365–376
7. Foster, I., Kesselman, C., eds.: The Grid: Blueprint for a New Computing Infrastructure. Morgan Kaufmann Publishers (1998)
8. Foster, I.: What is the grid? a three point checklist. Technical report, Argonne National Laboratory & University of Chicago (2002)

9. Equator/MIAS: Grid based medical devices for everyday health. http://www.gridoutreach.org.uk/docs/pilots/meddev.htm (2003)
10. De Roure, D., Baker, M., Jennings, N., Shadbolt, N.: The evolution of the grid. In Berman, F., Hey, A., Fox, G., eds.: Grid Computing: Making The Global Infrastructure a Reality. John Wiley & Sons (2003) 65 – 100
11. Raman, B., Agarwal, S., Chen, Y., Caesar, M., Cui, W., Johansson, P., Lai, K., Lavian, T., Machiraju, S., Mao, Z., Porter, G., Roscoe, T., Seshadri, M., Shih, J., Sklower, K., Subramanian, L., Suzuki, T., Zhuang, S., Joseph, A., Katz, R., Stoica, I.: The sahara model for service composition across multiple providers. In Mattern, F., Naghshineh, M., eds.: Pervasive Computing, First International Conference, Pervasive 2002, Springer (2002) 1–15
12. Rodden, T., Benford, S.: The evolution of buildings and implications for the design of ubiquitous domestic environments. In: Proceedings of the conference on Human factors in computing systems, ACM Press (2003) 9–16
13. Weiser, M., Brown, J.: Designing calm technology. In Denning, P., Metcalfe, R., eds.: Beyond Calculation – The Next Fifty Years of Computing. Copernicus Books (1997)
14. Boddupalli, P., Al-Bin-Ali, F., Davies, N., Friday, A., Storz, O., Wu, M.: A case for payments in ubiquitous computing environments. (Submitted for publication)
15. Foster, I., Kesselman, C., Tuecke, S.: The anatomy of the grid: Enabling scalable virtual organizations. International Journal of Supercomputer Applications **15** (2001)
16. Foster, I., Kesselman, C., Nick, J., Tuecke, S.: The physiology of the grid: An open grid services architecture for distributed systems integration. Technical report, Open Grid Service Infrastructure WG, Global Grid Forum (2002)
17. GLOBUS – The Globus Project: The globus toolkit. http://www.globus.org/toolkit/ (2003)
18. Grimshaw, A., Wulf, W., French, J., Weaver, A., Reynolds, P.: Legion: The next logical step toward a nationwide computer. Technical report, University of Virginia (1994)
19. Avaki: Avaki compute grid. http://www.avaki.com/products/acg.html (2003)
20. GGF: The global grid forum. http://www.gridforum.org/ (2003)
21. Tuecke, S., Czajkowski, K., Foster, I., Frey, J., Graham, S., Kesselman, C., Snelling, D., Vanderbilt, P.: Open grid services infrastructure (ogsi). http://www.gridforum.org/ogsi-wg/drafts/draft-ggf-ogsi-gridservice-23_2003-02-17.pdf (2002)
22. Benford, S., Crout, N., Crowe, J., Egglestone, S., Foster, M., Greenhalgh, C., Hampshire, A., Hayes-Gill, B., Humble, J., Irune, A., Laybourn-Parry, J., Palethorpe, B., Reid, T., Sumner, M.: escience from the antartic to the grid, Nottingham, U.K. (2003)
23. The UK Database Task Force: Open grid services architecture data access and integration (ogsa-dai). http://www.ogsadai.org.uk/ (2003)
24. Albinali, F., Boddupalli, P., Davies, N., Friday, A.: Correlating sensors and activities in an intelligent environment: A logistic regression approach. In: To appear European Symposium on Ambient Intelligence, Eindhoven, The Netherlands (2003)