Toward a Standard Ubiquitous Computing Framework

Martin Modahl [*]
Marsdorferstrasse 36
50858 Köln Germany
mmodahl@gmail.com

Bikash Agarwalla
Gregory Abowd
Umakishore Ramachandran
College of Computing
Georgia Institute of Technology
801 Atlantic Drive NW
Atlanta, GA 30332-0280
[bikash|abowd|rama]
@cc.gatech.edu

T. Scott Saponas [†]
Dept of Computer Science
University of Washington
Box 352350
Seattle, WA 98195-2350
ssaponas@cs.washington.edu

ABSTRACT

This paper surveys a variety of subsystems designed to be the building blocks from which sophisticated infrastructures for ubiquitous computing are assembled. Our experience shows that many of these building blocks fit neatly into one of five categories, each containing functionally-equivalent components. Effectively identifying the best-fit "lego pieces", which in turn determines the composite functionality of the resulting infrastructure, is critical. The selection process, however, is impeded by the lack of convention for labeling these classes of building blocks. The lack of clarity with respect to what ready-made subsystems are available within each class often results in naive re-implementation of ready-made components, monolithic and clumsy implementations, and implementations that impose non-standard interfaces onto the applications above. This paper explores each class of subsystems in light of the experience gained over two years of active development of both ubiquitous computing applications and software infrastructures for their deployment.

Categories and Subject Descriptors

C.2.2 [**Computer-Communication Networks**]: Networking Protocols—*Protocol Architecture (OSI Model)*; D.2.1 [**Software Engineering**]: Requirements/Specifications—*Methodologies*; D.2.11 [**Software Engineering**]: Software Architectures—*Patterns*

Keywords

ubiquitous infrastructure, infrastructure stack

---

[*]Martin completed all work while at Georgia Tech.

[†]Scott completed all work while at Georgia Tech.

1. INTRODUCTION

A ubiquitous computing (UbiComp) environment entails (1) an extensive and complex computer architecture deployment, (2) sophisticated conduits for dynamic data flow and control across this architecture, and (3) a judicious external interface facilitating user interaction with the system. A ubiquitous infrastructure must act very much like a telephone patch board i.e., a software layer which various assets are "hooked" into and where logical and physical connections between these assets are instantiated. Assets, in the broadest sense, refer to physical devices, bandwidth, processors, as well as services made available through the patch board. This analogy suggests a standard and uniform, but not necessarily central, interface. It is only through this middleware layer that the vast array of disparate assets can be brought together to create the type of rich ubiquitous applications we are excited about.

A necessary precondition for deployment of a UbiComp application is a distributed system composed of (1) computational resources including network-connected sensors and actuators, and (2) a mechanism to generalize patterns of interaction with these resources. The reusable corpora of middleware can speed the second step, but in our own experience, we have often overextended or mis-used existing UbiComp middleware subsystems. A standard language for describing middleware subsystems or a simple taxonomy for classifying subsystems could aid in the identification of potential candidates for a middleware deployment.

Different distributed applications, each using middleware for intra-application communication, have very diverse needs. Similarly, an individual distributed application might, at times, require very different things of middleware subsystems. In these cases, multiple, complementary middleware subsystems might be used simultaneously across a single computer architecture deployment to provide the most effective communication tools to all applications. Here the middleware taxonomy must also provide aid in choosing what multiple middleware subsystems might be deployed simultaneously to the greatest effect.

To arrive at a reasonable taxonomy describing UbiComp infrastructure and to understand the possible infrastructure interactions, we begin by examining some current and potential UbiComp infrastructures such as UPnP, MediaBroker, and the GRID. By partitioning these varied middleware subsystems into categories/ equivalence classes, it becomes possible to reason about entire infrastructures piecemeal. Thus a whole infrastructure may be evaluated

in terms of its constituent parts, streamlining the evaluation process and the infrastructure's subsequent refinement.

In addition, standard subsystem classes can improve development of standard intra-infrastructure, subsystem-to-subsystem interaction patterns. With a clearer picture of what infrastructure capabilities exist and where they are located, developers are better positioned to leverage these capabilities in their applications. Once the application requirements have been identified, developers may be able to pick existing infrastructure pieces to develop the application quickly and effectively.

This paper proposes a five-class infrastructure taxonomy based on the orthogonal functionalities of most commonly occurring subsystems. These include *Registration and Discovery*, *Service and Subscription*, *Data Storage and Streaming*, *Computation Sharing* and *Context Management*.

Figure 1, appearing later in the paper, shows a stack corresponding to top to bottom interaction between the equivalence classes. The user application is provided the APIs for interacting with Services, Streams and Storage Locations, and Process Migration/ Distribution. Context Management Services and Ontologies are also common among all applications.

This paper continues in Section 2 discussing the imagined computer architecture deployment for UbiComp as well as some initial application deployment experience using middleware to bridge between application and hardware deployment. Section 3 details the five taxonomy classes. We conclude in Section 4.

2. EXAMINING EXISTING INFRASTRUCTURE COMPONENTS

The intrinsic nature of common UbiComp applications builds on connecting distributed nodes into a over-arching application. This application might distribute itself onto multiple computers as a common Internet messaging client is deployed on multiple computers throughout the world, while each individual client connects into a larger whole that is a global messaging application. An application might also run in a single location but request and access resources throughout an environment the same way a cable set-top box might request resources from the cable company and display those resources on, what in the future might be, an autonomous display.

A common infrastructure built to support UbiComp applications must then provide the ability for applications to access computation and data resources throughout the environment. Previously we discussed UbiComp infrastructure as the combination of deployed computer hardware and the middleware stitching it together. Before we discuss the middleware subsystems running across the hardware, we will discuss the hardware deployment paradigm we are working from and our previous experiences building applications on top of hardware deployed according to this paradigm.

2.1 A Deployment Paradigm

Striving toward the deployment of UbiComp applications in smart spaces, we have been developing both applications and infrastructures with a simple paradigm in mind. This paradigm is constructed of simple, modular, moderately-capable computing devices being connected directly to sensors or sinks throughout a space. In this paradigm each computer can proxy access to its directly connected devices onto network-accessible interfaces. Transient devices and stationary devices are connected together abstracting over connection media to form connected graphs of computational devices. Each device then is a node in the graph.

Each of these distributed nodes will provide simple building block resources that applications might query and use. This way, an application might require a camera stream or a speaker, query for devices containing that service in the area requested, then request and use the resource. Each of these resources as well as the infrastructures exposing these resources to the outside world are managed as modular building blocks.

2.2 Development Case Study

Our software intercom is an example of an application built on hardware and middleware, deployed with the above paradigm in mind, in the Aware Home at Georgia Tech [17]. Using consumer electronics computers and our middleware—MediaBroker [20]— this software intercom allows users throughout a home to communicate through walls and floors as if they were in the same room.

In accordance with the paradigm described above, applications expose resources which other applications use. The user interacts with an intercom interface application which controls a set of modular audio applications deployed throughout the Aware Home. The user interface application projects a control-information resource which outputs current state information while the audio applications provide audio resources which stream captured audio. While running, The interface application queries the state of the audio resources and actuates end-to-end connections through its single control resource. In turn each audio application listens to the interface application's control resource for direction.

Our initial deployment, built entirely on MediaBroker, relied on one-way data streams to transfer audio information between audio inputs and outputs between audio access applications while using the same one-way data streams to distribute control information from the interface application to the audio applications. The nature of these one-way, relatively static, connections between controller and controlled made the deployment of the intercom much less dynamic than anticipated.
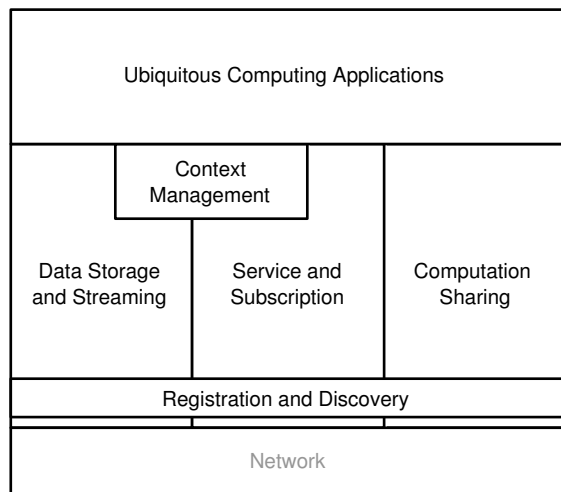
Recently, the intercom has been redesigned to use MediaBroker as an infrastructure for low latency audio data streams and UPnP for distributing discrete control messages between a set of control applications and the audio applications they control.

In this case, taking advantage of UPnP for low-bandwidth message passing allows us to apply MediaBroker to the problem it better solves —moving audio data from sources to sinks according to data requests— while offloading control message handling to UPnP. UPnP has better facility for creating the interface application's control resource and for allowing the many audio access applications to better listen to multiple interface applications at once.

Using the strength of each subsystem and the scaffolding each offers, the job of programming the audio application and the control application is greatly simplified. Using the standard interfaces provided by UPnP and MediaBroker cleanly, the programmer can concentrate on the individual concerns of each application and not how to translate those concerns into something the middleware can do.With a well-defined taxonomy with which to understand UbiComp middleware subsystems, implementation of this intercom would begin with brief modeling of all intra-application and inter-application interactions based on the capabilities of each middleware subsystem categories' abilities.

3. SUBSYSTEM CATEGORY OVERVIEW

In the brief case study above, system deployment and application development were made simpler through leveraging an existing technology in the context of a novel problem. This anecdotal evidence is in fact exemplary of the benefits derived through clean separation of subsystems. As we have seen, the delineation and choosing of subsystems along functional boundaries enables us to

```
┌─────────────────────────────────────────────┐
│         Ubiquitous Computing Applications     │
├──────────┬────────────────────┬───────────────┤
│          │      Context        │               │
│          │    Management       │               │
│ Data Storage ├─────────┬────────┤ Computation   │
│ and Streaming│ Service │        │   Sharing     │
│          │   and   │        │               │
│          │ Subscription│     │               │
├──────────┴─────────┴────────┴───────────────┤
│         Registration and Discovery            │
├───────────────────────────────────────────────┤
│                  Network                       │
└───────────────────────────────────────────────┘
```

Figure 1: OSI Application Layer with added UbiComp Infrastructure Component Classes

better reason about the infrastructure as a whole as well as to improve the quality of a specific application. Each of these existing subsystems filled a specific role in the development of the application and in the larger infrastructure deployment.

Similar in functionality to UPnP [8], most WebServices implementations provide the ability to make remote calls to procedure located in other applications on other hosts. We argue, then, that UPnP and WebServices provide interchangeable functionality, which means that they belong to the same category of subsystems, namely Service and Subscription, and might be fairly compared to each other when selecting the best-fit subsystem for the job at hand. In turn, their current limitations may serve as a driver for development of a more robust replacement.

In turn, the high-bandwidth, low-latency streaming facilities offered by MediaBroker is what earns it its place in a different category of subsystems, namely Data Storage and Streaming. Other subsystems exist to accomplish similar stream-oriented tasks: GnuStream [16] and D-Stampede [1], while still other provide shared network accessible data stores: Coda [24] and T-Spaces [18].

3.1 Registration and Discovery

Communication between applications running on top of the distributed nodes of a UbiComp infrastructure begins with discovery of available resources. Once different applications connect, they can share descriptions of their capabilities and begin to collaborate on higher order tasks.

Myriad languages, protocols and frameworks exist for the discovery of hosts on the network, applications running on hosts, and services and resources provided by those applications. Common to the WebServices and UPnP world are protocols like SSDP [14], UDDI [3], and WSDL [15]. These allow robust and systematic description and discovery of services. Jini provides significant functionality in the area of service discovery as well as actual resource migration once it is discovered. Many lightweight services like DNS [19] and LDAP provide a more static registration mechanism suitable for resources with considerably longer lifetimes. Specialized facilities like BSDP, which is tied to Bluetooth protocol stack, are also popular.

Evaluating registration and discovery subsystems based on their capabilities must take into account efficiency and extensibility. Some

subsystems will require strict and complex registration as well as manual discovery and iterative search of what is available. Other subsystems will provide functionality to automate much of this.

3.2 Service and Subscription

Service and subscription encompasses the passing of discrete messages to accomplish distributed service registration and access. For example, an application might offer a resource capable of converting Unites States zip codes to map images [27]. Applications can then access this resource to accomplish this. Applications might also want to know when some event happens and they might subscribe to be notified of some change. Both service and subscription rely on discrete messages being reliably passed between applications.

According the paradigm discussed above, UbiComp applications will consist of many smaller parts that will require intra-application communication for synchronization. The discrete information exchange provided by a service and subscription subsystem allows distributed application entities to keep state between themselves and provides the necessary abstractions for setting state on adjacent nodes and being updated with their states when other nodes' states change.

The service and subscription component must provide an extensible application programming interface. The component must also be able to scale with respect to the number of applications using the subsystem and the size of data being transfered between applications. The speed with which remote services are called and with which subscription updates are propagated is also integral to the evaluation.

The service and subscription subsystem we have discussed previously in this paper, UPnP, is really only mediocre at message passing and has limited functionality allowing extensibility. UPnP's scalability is also limited to a single network subnet without resorting to complex bridging technology. Context Toolkit [23] provides a service and subscription component as well as some basic context reasoning ability but it has many of the same limitations of UPnP in its inefficiency. Webservice implementations vary greatly, but many provide very efficient messaging and fairly extensible application programing interfaces.

3.3 Data Storage and Streaming

Beyond the simple, extensible message passing described above, many UbiComp applications, especially those geared toward multimedia capture and access, require significant, structured data transfer. We define the data sharing and streaming component as a component capable of handling significant structured data movement between distributed nodes. This data movement might involve streaming multimedia from a media provider to a media consumer or it might encompass more complex requirements for data archival and retrieval.

Streaming middleware subsystems like MediaBroker [20], GnuStream [16], Indiva [22] and D-Stampede [1] allow distributed nodes to ship streams of data to each other, while other subsystems like Serverless File Systems [2] and Coda [24] provide distributed access to common file systems. Infrastructures like T-Spaces [18] or JavaSpaces [26] provide more generic shareable network memory. All of these classes provide access to large amounts of data in a structured manner.

Measuring the performance of a data storage and streaming subsystem might be difficult if we only take into account efficiency at providing quality of service. Many data storage and streaming subsystems provide varied types of functionality including different delivery or data set management schemes.

3.4 Computation Sharing

Computation sharing allows a running application to make use of remote computational resources on demand. For example, a surveillance application running on a sensor network may decide to offload complicated vision tracking components onto more capable high performance computing resources (HPC) strategically placed throughout the ubiquitous infrastructure. Furthermore, in order to accommodate mobile clients to deliver required quality of service, such computation must be allowed to move from one HPC resource to another. In order to support such unfettered, yet highly-specialized access to computing cycles, a mechanism for discovering unclaimed CPU time and dynamically scheduling computations on the HPC resources is essential.

Many tools for enabling distribution of compute-intensive tasks have emerged over the years. The scientific computing community is largely responsible for this variety. Regrettably, the focus has often been on batch processing, which alone is insufficient to support rich ubiquitous computing applications we envision. Only recently, the grid computing [11, 12] community has started to focus on applications that require capabilities beyond batch processing. Grid computing allows controlled sharing of computation resources across many different physical organizations in order to deliver a quality of service. Open Grid Service Infrastructure(OGSI) [11] has extended web services by providing standard mechanism for creating, naming and discovering grid services, providing virtualization through location transparency, and supporting integration with underlying native platform facilities. The Globus Toolkit [10] built on OGSI model, provides the middleware support needed to build grid computing applications.

Grid services provide support for service and subscription as well as computation sharing. Computation sharing is provided through the resource management systems [21] built for the grid. Condor [4], nimrod-g [5] and legion [7] are some of the projects that have emphasized building meta-schedulers for running batch jobs on the grid. Grid Resource Allocation and Management (GRAM) framework in Globus provides remote access to individual resources and also integrates with the native facility on each resource. In summary, computation sharing is provided through the resource management system of the grid. Moreover, since grid services are built on top of web services, service and subscription component of our taxonomy is provided through appropriately built grid services.

MPI [9], PVM [13] are some of the other tools that can be used for resource sharing within a cluster. These tools provide libraries for building scientific computing applications and running them on the cluster. However, the capabilities of these tools doesn't extend very well for streaming applications. Moreover, they don't provide capabilities for dynamically asking additional resources or dynamic joining of computational entities.

Facilities for computation sharing must be dynamic enough to provide for application, thread or even function migration across multiple distributed entities. Applications might want to dynamically move through the environment in order to physically follow a user. Applications might also simply want to spread computationally intensive processes to more capable elements in the environment. The tools for computation sharing should also support dynamism in terms of the resources being allocated, the computation being performed as well as the mobility of the user.

3.5 Context Management

A context management subsystem defines a common language to describe the context of an environment and also defines a language to discuss described context state. This middleware subsystem is built on both the service and subscription components as well as the data sharing and streaming components as it will commonly differentiate streaming sensor data into discrete values. While this differentiation will happen at the ubiquitous application level, the differentiation process can be represented inside the context state.

Much of UbiComp relies on having access to a consistent view of the world so that applications might correctly actuate themselves according to the environment and those in the environment. This common context management subsystem component allows applications to keep current on the state of the world by communicating with each other in a language common to all possible applications in the space.

Important in the realm of context management, synonym and unique naming management have been studied fairly thoroughly by multiple projects under the SemanticWeb [6] umbrella and also by projects like OntoWeb [25] and Context ToolKit [23].

4. CONCLUSION

In this paper we have presented a comprehensive UbiComp middleware taxonomy to ease discussion and facilitate the process of building both ubiquitous middleware subsystems as well as ubiquitous applications. We also discussed where in our taxonomy existing middleware subsystems fit. In our experience, the sampling of industrial and academic experience and the combination of knowledge across groups allows better middleware construction and this paper hopes to further these interactions through the creation of a common language for middleware discussion. We are not entirely sure about the feasibility of creating complex infrastructures by picking and choosing subsystems from each taxonomy category, but we look to examine it. Much research is still to be done at the intersections of these subsystem categories to explain the gaps that exist and also to examine their interactions. Research can also be directed to survey entire middleware categories; learning from the state of the art and designing and building new class-complete infrastructures.

4.1 Acknowledgments

5. REFERENCES

[1] S. Adhikari, A. Paul, and U. Ramachandran. D-stampede: Distributed programming system for ubiquitous computing. In *Proceedings of the 22nd International Conference on Distributed Computing Systems(ICDCS)*, Vienna, July 2002.

[2] T. Anderson, M. Dahlin, J. Neefe, D. Pat-terson, D. Roselli, and R. Wang. Serverless network file systems. In *Proceedings of the 15th Symposium on Operating System Principles*, pages 109–126, Copper Mountain Resort, Colorado, December 1995.

[3] Bellwood and et al. UDDI version 2.03 data structure reference, July 2002.

[4] R. Boer. Resource management in the condor system. Master's thesis, Delft University of Technology, 1996.

[5] R. Buyya, D. Abramson, and J. Giddy. Nimrod/g: An architecture for a resource management and scheduling system in a global computational grid. *High Performance Computing (HPC) ASIA*, 2000.

[6] D. Calvanese, G. D. Giacomo, and M. Lenzerini. A framework for ontology integration. In *The First Semantic*

Web Working Symposium, pages 303–316, 2001.

[7] S. J. Chapin, D. Katramatos, J. Karpovich, and A. S. Grimshaw. The Legion resource management system. In D. G. Feitelson and L. Rudolph, editors, *Job Scheduling Strategies for Parallel Processing*, pages 162–178. Springer Verlag, 1999.

[8] Contributing Members of the UPnP(TM) Forum. Welcome to the UPnP(TM) Forum, 2003. http://www.upnp.org (March 1st, 2004).

[9] M. P. I. Forum. MPI: A message-passing interface standard. Technical Report UT-CS-94-230, 1994.

[10] I. Foster and C. Kesselman. Globus: A metacomputing infrastructure toolkit. *The International Journal of Supercomputer Applications and High Performance Computing*, 11(2):115–128, Summer 1997.

[11] I. Foster, C. Kesselman, J. Nick, and S. Tuecke. The physiology of the grid: An open grid services architecture for distributed systems integration. Technical report, 2002.

[12] I. Foster, C. Kesselman, and S. Tuecke. The anatomy of the Grid: Enabling scalable virtual organizations. *Lecture Notes in Computer Science*, 2150, 2001.

[13] A. Geist, A. Beguelin, J. Dongarra, W. Jiang, R. Manchek, and V. Sunderam. *PVM Parallel Virtual Machine, A User's Guide and Tutorial for Networked Parallel Computing*. MIT Press, Cambridge, Mass., 1994.

[14] Y. Y. Goland, T. Cai, P. Leach, and Y. Gu. IETF Internet Draft: Simple service discovery protocol/1.0, Oct. 1999.

[15] M. Gudgin, A. Lewis, and J. Schlimmer. W3C Working Draft: Web services description language (wsdl) version 2.0 part 2: Message exchange patterns, Mar. 2004.

[16] X. Jiang, Y. Dong, D. Xu, and B. Bhargava. Gnustream: a p2p media streaming prototype. In *Proceedings of IEEE Internation Conference on Multimedia and Expo*, July 2003.

[17] C. D. Kidd, R. Orr, G. D. Abowd, C. G. Atkeson, I. A. Essa, B. MacIntyre, E. D. Mynatt, T. Starner, and W. Newstetter. The aware home: A living laboratory for ubiquitous computing research. In *Cooperative Buildings*, pages 191–198, 1999.

[18] T. J. Lehman, S. W. McLaughry, and P. Wycko. T Spaces: The next wave. In *HICSS*, 1999.

[19] P. V. Mockapetris. RFC 1101: DNS encoding of network names and other types, Apr. 1989.

[20] M. Modahl, I. Bagrak, M. Wolenetz, P. Hutto, and U. Ramachandran. MediaBroker: An Architecture for Pervasive Computing. In *Proceedings of the 2nd IEEE International Conference on Pervasive Computing and Communications*, Mar. 2004.

[21] J. Nabrzyski, J. M. Schopf, and J. Weglarz. *Grid Resource Management: State of the Art and Future Trends*. Kluwer Academic Publishers, 2003.

[22] W. T. Ooi, P. Pletcher, and L. A. Rowe. INDIVA: Middleware for managing a distributed media environment. (BMRC Technical Note), 2002.

[23] D. Salber, A. K. Dey, and G. D. Abowd. The Context Toolkit: Aiding the development of context-enabled applications. In *CHI*, pages 434–441, 1999.

[24] M. Satyanarayanan, J. J. Kistler, P. Kumar, M. E. Okasaki, E. H. Siegel, and D. C. Steere. Coda: A highly available file system for a distributed workstation environment. *IEEE Transactions on Computers*, 39(4):447–459, 1990.

[25] P. Spyns, D. Oberle, R. Volz, J. Zheng, M. Jarrar, Y. Sure, R. Studer, and R. Meersman. Ontoweb - a semantic web community portal. In *Fourth International Conference on Practical Aspects of Knowledge Management (PAKM)*, Dec. 2002.

[26] Sun Microsystems. Javaspace specification, 1998.

[27] Terraserver.Com, Inc. TerraServer.com, 2004. http://www.terraserver.com (March 1st, 2004).