Supporting Nomadic Healers

By Rajive Bagrodia, Mario Gerla, Songwu Lu, Richard Meyer, Daniel J. Valentino, and Lixia Zhang

Abstract

In the traditional Internet application model, an end user interacts directly with the application server. While this is a simple model of interaction between end user clients and application servers, it is being severely challenged by the growing number of mobile wireless users and the diversity of client devices. Supporting wireless and mobile access to the Internet for healthcare providers is one instance of a universal trend that data access is increasingly becoming wireless and mobile. The user mobility raises a number of issues that include user authentication, communication security, providing adequate application-level QoS, and network scalability. This paper suggests that an appropriate solution to this problem lies in the design and deployment of a middleware infrastructure to help bridge the gap between application servers and large numbers of nomadic end users, and between the application requirements and network conditions. A number of challenges to the design of such a middleware infrastructure are discussed.

1. Introduction

We envision that within the next few years mobile and wireless access to the Internet will very likely become the norm, rather than the exception it is today, and *anytime*, *anywhere*, on *any platform* access to multimedia data will be taken for granted. Market studies, such as those from the Yankee group, report that nearly 40 million workers in the US are already using portable computers, and laptops and handheld PCs are the fastest growing segment of the computer industry. 60% of mobile computer users want wireless mobile access to their information, and current estimates of the potential market for wireless data range as high as $37.5 billion in revenues in 2002 for wireless Internet applications alone (Killen & Associates). These advances in technology creates a new challenge for the information infrastructure to provide ubiquitous information access and sharing to enable mobile users to become first class network citizen, able to access and manipulate information online while roaming seamlessly.

In the traditional Internet application model, each client contacts application servers directly to fetch needed data. This simple model imposes a burden on data servers, clients, and the underlying network. Data servers, in addition to performing their main task of collecting and maintaining application data, must also take responsibility for data dissemination and access control. They interact with a host of clients, who may have a variety of different capacities along their access routes or in their display format, and handle all phases of the data delivery process from user authentication to congestion control. Many servers have been developed purely as an application database, and are ill suited for handling large numbers of heterogeneous users.

Wireless mobile access to the network brings new challenges to this simple model of direct communication channels. When a mobile host moves from place to place, the movement may lead to dramatic variations in channel quality. Furthermore, the mobility raises new issues in user authentication and communication security as users change locations and IP addresses during an application session. A central tenet in this vision is that users are more mobile than the devices they carry. The infrastructure must evolve to allow users seamless access to data from any device on a variety of networks, from wearable devices using CDPD or Bluetooth, to high resolution graphic machines tied to Internet2. An important challenge is to develop a middleware infrastructure to: a) help bridge the gap between application servers and nomadic clients as they migrate active sessions among multiple devices; b) balance application requirements and network conditions; and c) enhance the scalability of the network. In iMASH, the middleware servers are treated as first class network citizens, visible entities that lie between the mobile clients and the application servers.

The iMASH project, a joint collaboration between the Computer Science Department and the Medical School, is developing the framework of such an adaptive middleware infrastructure, initially for healthcare workers. Its objective is to provide the capability for real-time, multimedia communication, such that a physician can access, on the move, the patients record and related information, and can migrate ongoing application sessions seamlessly to multiple platforms that may range from a high performance diagnostic workstation in the physician's office to hand held PDAs in the examination room.In this short paper we present a number of challenges in developing this middleware framework. The next section describes the health care application targeted by this project and the requirements that it imposes on the computing and communication infrastructure. Section 3 describes the system architecture for iMASH and the middleware services it must provide Section 4 outlines the major design challenges faced in the iMASH project. Section 5 describes preceding related work in the area.

2. Application Requirements

The patients record system in a typical hospital consists of diverse types of information stored in multiple databases that may potentially include digital images from a radiology information system, pathology images, laboratory test results and reports, electronic patient data in a hospital information system, and related information derived from the original data. As more data is acquired in digital format, the need for ubiquitous access to the data has also grown tremendously. Access to the electronic medical record is required in many more locations and situations than can be met using desktop computers and traditionally wired devices. A new information system architecture is needed that utilizes middleware infrastructure to provide a number of sophisticated, computer-based services. The middleware services and protocols enable data to be displayed and manipulated on a variety of devices ranging from hand-held computers to high-resolution diagnostic workstations. This includes access to electronic patient records consisting of images, reports, demographic information, and other medical databases, and the distribution of medical images and records over the Internet for telemedicine and for long-distance medical education. These capabilities make it possible to better track patient care and to transmit the appropriate patient records to the appropriate health care providers. The providers include the primary care physician, sub-specialist physicians consulting on the case, and even to the patients themselves so that they can better understand their medical situation.

Consider the following scenario that exemplifies the dramatic manner in which iMASH may alter patient care. A neurologist receives a page while he is driving to work. She answers the page and is asked to consult on a patient, Mr. Jones, with suspected neurological disorders. She tells the referring primary care physician to transfer Mr. Jones' electronic medical record to the hospital. The hospital system receives the complete record, and begins transmitting the clinically relevant information to the physician's "Mobile Physician Assistant" (MPA). Using text-to-voice technology, the MPA begins reading the key information to the neurologist. She agrees that Mr. Jones needs further neurological evaluation, and dictates some orders (tests) through the MPA, which then places the orders with the hospital system. The system sends the orders to the appropriate sub-specialists, for verification (e.g., imaging), and each of them uses an MPA to review the order and confirm it. The system notifies the schedulers to call Mr. Jones and schedule the procedures. The system will notify the neurologist when Mr. Jones has had the tests performed, automatically transmit them to the neurologist's location.

Upon arriving in the hospital, the neurologist goes to the clinic to meet the first patient, Mr. Smith, scheduled for that day. She reviews Mr. Smith's electronic medical record (EMR) on his MPA, which the system has previously found and transmitted based upon the schedule. The EMR is customized for his neuro-oncology evaluation and presented in a way that is appropriate for the display on the MPA that he is using. The Adaptive Viewing Protocol automatically displays pertinent demographics, details of the primary neoplasm, including the primary site and metastatic sites and stage. She can quickly download the current treatment regimen, including cycle and day from baseline, and with links to full descriptions of the each treatment. Based upon his evaluation of this data, she decides to look at the old images for this patient. Her MPA does not provide a good enough display, so she moves to the ultra-high resolution diagnostic workstation that is centrally located between exam rooms. After she identifies herself via fingerprinting, the workstation display is activated and more detailed information is automatically presented. She reviews a multimedia presentation of images, notes, and a graphical display of measurements that were compiled by a neuro-radiologist, and based upon that information, decides that Mr. Smith is improving, and confidently reduces his chemotherapy by a small amount. She enters the change to the medication via her MPA, and walks into the exam room to see the next patient. At the same time, the next patient's EMR is automatically transferred to the MPA, while Mr. Smith is sent a page and email, reminding him to check his medical website for an update on his clinical condition. The neurologist stops at the clinical station to leave a video-mail for her colleague to let him know that Mr. Jones will be coming to the hospital for an evaluation, and to alert him that his particular expertise might be needed.

Realization of this scenario requires appropriately leveraging technological innovations in the commercial sector with innovative use of networking and middleware capabilities. The hardware devices referred to in the scenario are either available or likely to be available within the next two years. The capabilities needed to support such a nomadic healer include the following, among many other things.

- *Location tracking*. The system must track the nomad's location, in order to adapt to changing network conditions, to provide QoS by pre-allocating resources, to provide seamless session transfer as network boundaries are crossed.
- *Adaptive displays*. As doctors move from room to room, they will need to make use of the display equipment available to them. The system must be able to adapt the data to the display characteristics of many devices.
- *Session handoff*. In order to move seamlessly between devices, copies of session state must be maintained in a lightweight format for easy transfer. The framework for session handoff is conceptually similar to [1].
- *Security* and *authentication*. To protect the privacy of patients and comply with newly enacted legislation, the system must be secure at all levels.
- *Service discovery*. As devices are activated or move into new subnetworks, they must quickly and efficiently

discover the services they require.
- *Wide area data caching* and *consistency management*. A patient's data may be accessed several times in quick succession. To improve QoS and reduce network load, that data may be cached locally in middleware, but that raises issues of data consistency, as multiple local copies must be kept up to date.
- *Scalability*. A central challenge is to design this capability such that the system is able to scale in the number of users, the number of devices, and the physical area covered without significant performance degradation.

In the rest of this paper, we explore the challenges in developing such a capability based on the notion of *distributed middleware servers* to accomplish this objective.

3. Middleware and System Architecture

In addition to software and protocol support, the iMASH project will deploy a hardware/software infrastructure, a la Wireless Andrew [2], adapted to the healthcare domain. The planned hardware architecture consists of
- A set of application servers connected to a high bandwidth wired network.
- A collection of wired devices that includes office workstations, clinic stations, imaging equipment, etc.
- High bandwidth wireless access within the hospital, augmented by one or more lower bandwidth wide area wireless systems.
- A set of strategically deployed ,dedicated middleware servers.

The basis of the iMASH architecture is the distributed set of *middleware servers*. These servers, when viewed at the application level, interconnect all applications and clients, and are treated as first-class network citizens. That is, they are visible entities at the application level, rather than hidden boxes between clients and applications to provide only performance enhancement. Requests from clients are *explicitly* addressed to nearby middleware servers, and middleware servers are the source for all services for clients, simplifying the problem of service discovery, as well as the sole client for application servers, considerably simplifying their role. The deployment of middleware servers between the clients and servers provides scalability in several ways. Application servers need not be concerned with a large number of end users, but instead can concentrate on collecting and processing data. A client device need not be concerned with contacting potentially large numbers of servers to gather data, but instead contacts only the local middleware server for all services. Middleware takes responsibility for getting the data from the right servers, and makes necessary conversion to fit the clients' needs. Middleware can also help scaling by temporary caching, fetching each piece of data only once but serving multiple requests for the same data, for example, as a physician moves locally between a clinic workstation and an MPA. Middleware servers provide the following functionality:
- *User authentication and profiling*. Middleware boxes are trusted agents. They authenticate end users on behalf of the application.
- *Information discovery/data caching*. The middleware server acts as a proxy on behalf of the client to retrieve data from the application server. It is expected that several doctors or nurses may need to access records for a single patient within a short time span. As with WWW caching, the middleware can significantly reduce network load as well as response time by caching a local copy of the data.
- *Presentation conversion*. Middleware servers fetch data based on user requests (or pre-fetch data based on prediction of user's near-future need), and perform conversion as needed. Thus middleware servers must be user-aware (need user profile), network-aware (need network condition), data-aware (need codecs/encryption keys), and device-aware (need knowledge of different boxes).

The preceding services wold be required even in a wired environment. Mobility brings additional requirements to middleware service:
- *Location tracking*. To maintain an active application session as a user moves, the server tracks the user's location, and adjusts to local network conditions by pre-allocating bandwidth and other resources along the user's path. The move may also lead to a change of middleware servers, in which case the session must be forwarded to the nearest middleware server.
- *Session data caching*. When a user moves an on-going application session from one device to another, middleware servers act as a "home" for the application state (including active connections, cached data, etc.) to facilitate migration between devices.
- *User re-authentication*. When a user changes devices, or spawns a new branch of a session to a new device, the middleware server authenticates the user on the new device, and arranges to send data to the new address, with corresponding changes in presentation.

Because of the last, an example of how users are more mobile than the devices they carry, moving easily between devices and networks, users' application sessions are no longer associated with a specific *home* computer. Users' profiles and session data are therefore stored on the network, and retrieved onto a device during the user authentication process.

4. Design Challenges

A number of significant design challenges are introduced by the proposed use of middleware described in the previous section, including: *middleware server discovery, data consistency, security, QoS-driven protocols, adaptive user displays,* and *scalability*.

Server Discovery:

For a new/mobile client, server discovery can use a broadcast/listen model: each middleware server broadcasts its existence within a limited scope, and a mobile client listens to the server(s)' broadcast announcements and determines an appropriate server to which it subsequently sends a service request. The server broadcast approach is more efficient than the popular client announcement approach (e.g. in MASH) in our target mobile networking environment where the number of clients is significantly larger than that of servers and these clients may move frequently. Neighboring servers may have overlapping broadcast scopes, which are motivated by two performance criteria:

- seamless session handoffs and continuous data delivery during session handoffs; and
- load balancing (in a limited way) among neighboring servers.

iMASH's server discovery is distinquished from *service* discovery in systems such as Jini [3] and ICEBERG [4] because the services in iMASH are well known, and managed by middleware, so only the server need be discovered.

Consistency Management

Because a patient's data may be both accessed and modified several times in a short period, for example if one or more specialists consulting on the case enter notes into the record, consistency must be maintained between cached copies of the data. Weak data consistency can be achieved in the distributed middleware infrastructure via an approach similar to the consistency management schemes that have been developed for distributed shared memory computers [5] and replicated file systems [6]. It is possible that different data items in the patient record system use different forms of consistency. A possible directory based scheme is to require the application server (or a *home* server in case of a distributed application server) to keep a record of which middleware servers have 'cached' copies of a specific record. When the data copy at a middleware server is updated by a mobile client (e.g. a doctor may write new prescriptions for a patient, which will be input into the patient's record) and subsequently the middleware server updates the application server, the application server notifies the relevant middleware servers that their data copies are no longer valid. Each middleware server decides whether and when to fetch an updated data copy depending on its current demand or delete the old copy if no one has accessed that data copy during a certain period of time. Overloading in each middleware server is prevented by terminating server broadcast announcements when the service load reaches beyond a certain threshold on a middleware server.

Because the prior described system imposes a burden on the application servers, some of which may be legacy systems, another alternative is to put the burden of maintaining data consistency squarely on the middleware. A system similar to that employed by WWW proxies may be sufficiently effective and easy to deploy. HTTP servers provide the capability of checking whether data has been updated without downloading it. If the proxy's cached copy is up to date, it need not load a new copy. The middleware can either verify its cached data each time, or assign an expiration date after which the data must be refreshed or deleted.

Design of Adaptive Viewing Protocols

The design of adaptive displays whose contents adapt to device characteristics as well as user profiles and network conditions, is an important component of the middleware. Several approaches have been taken to providing adaptivity to varying network connectivity. These can be divided into three categories: 1) application-oriented, such as Odyssey [7] and the Rover Toolkit [8], where the middleware notifies the application of network conditions, 2) proxy-oriented, such as TACC [9], in which data is adapted at proxy nodes, and 3) protocol oriented, such as Protocol Boosters [10] and Transformer Tunnels [11], which try to make adaptivity transparent to the application.

Whereas adapting content to device characteristics and network conditions has been explored before, adaptation based on user profiles can be very useful in the context of specific applications. A promising approach in the medical context is the use of models of clinical processes to develop tailored viewing protocols that optimize the organization and presentation of multimedia data on the available viewing device. To demonstrate the relevance of this approach, we are developing object-oriented models (process models) of the clinical tasks performed by physicians and of the content used and created (data models), with a focus on the oncology domain. These models will be enhanced

- To use the process and data models to design innovative **adaptive viewing protocols** for the optimal organization and presentation of medical information, including images, text, audio, and related clinical documents used in neuro-oncology.
- To use the tailored viewing protocols to dynamically adapt the presentation of information on a **mobile display** used by a neuro-oncologist for the evaluation of patients with brain tumors.

The focus on this project is to design adaptive viewing protocols that will dynamically update the graphical user interface as the physician moves to different display devices, presenting the content and tools in a manner that is optimized for each device. Based upon the device characteristics and priority of the data that is available, the most appropriate viewing protocol will be automatically selected for the device. Within a protocol, the data and tools will be presented in a way that is appropriate for that device.

We have done preliminary work creating adaptive protocols based upon the direct input of neuroradiologists for studies such as a Brain Tumor MRI [12]. Based upon these preliminary results, we recognize the need to present complex data to physicians in a simplified manner. Using a workstation that we call the NeuroNavigator, we developed a protocol that integrates reports, images, and multimedia tools to quickly read a brain tumor case in a natural way following the high level goals of neuroradiologists. The protocols allow for automated presentation of MR comparisons, thus significantly speeding image analysis and facilitating report generation.

Security

To protect the privacy of patients, and to conform to a family of new legal requirements for electronic patient records, the iMASH project must provide security at every level of the hospital network. The goal is to restrict access to a patient's record to authorized personnel (and the patient herself), while not placing an overwhelming burden on hospital staff. The security requirements include:

- Third party (middleware) authentication. The direct communication model between application servers and clients requires a simple security model, wherein the application directly authenticates clients and controls access to the services and data. The introduction of middleware servers complicates the picture, because the application servers no longer interact directly with the clients: they must trust the middleware servers instead, to authenticate clients and deliver data. Introducing a third party in the process cannot be an enhancement to security, because each new box potentially introduces a new target for attack. To minimize the threat of such attack, one solution is to authenticate each middleware box for a finite time only. Whenever prolonged services are required, the middleware server must periodically re-authenticate themselves with the application servers.
- User identification and authorization. A system must be in place to allow doctors, nurses, patients, and hospital staff to identify themselves to the system, whether by username/password, fingerprints, or ID card/PIN. A framework should be designed to allow commercial solutions to be plugged in.
- Device identification. Devices must have access control profiles, as well as device characteristic profiles. For example, a display terminal in a public area may be restricted from displaying patients' financial data. Devices in secure areas may be allowed to maintain an idle session indefinitely, while a mobile device may require periodic re-authentication.
- Data access control.
- Logging of user/device activities.
- Protection from eavesdropping on both the wired and wireless networks.
- Re-authentication. As users transfer sessions from device to device, or as they move around, the middleware must re-authenticate them as transparently as possible.
- Proximity based or predictive authorization. If a doctor enters an examination room carrying an MPA, she may wish to quickly load images onto a wall-mounted HD display. The wall-mounted device should be loaded with the doctor's privileges automatically upon her entering the room with her MPA, so that she needn't laboriously log in as she enters each room. Preferably, the device would be controlled directly by the MPA, but a wall-mounted device is likely to be hardwired to the network and can most efficiently load data directly from middleware. One possibility is to use Bluetooth for local area session transfer and device control. Another possibility here is to do predictive authorization based on the doctor's schedule or mobility pattern.

Challenges in network adaptivity

Compared to wired networks, mobile wireless networks introduce two unique design issues: wireless channel dynamics and user mobility. The wireless medium is error-prone which causes the total effective channel capacity to change dynamically. The spatial contention that exists among simultaneously transmitting flows in a spatial locality, may also cause the effective bandwidth perceived by each user to change frequently. Both factors lead to significant resource dynamics in a wireless network environment. As she roams around the network, a mobile user may demand network resources from the new location that she intends to move into and the current location may reclaim the resources consumed by the same user; thus resource availability is location dependent and changes based upon user mobility. A fundamental problem in networking design for wireless mobile networks is that the wireless channel is a scarce and shared medium, exhibiting dynamics in both the time and spatial domains, and has to be managed effectively and adaptively.

In recent years, researchers have proposed numerous network design and management schemes for wireless mobile networks. In many of these proposals, adaptation has been identified as a key instrument to effectively cope with these dynamics. Many adaptive designs have appeared in the literature, including adaptive error control and recovery, adaptive resource reservation, adaptive mobility management, rate adaptation, wireless fair queuing, wireless TCP protocols, and adaptive video/audio codec, to name a few. The key idea behind these proposals is to provide effective and efficient adaptation in order to address wireless channel errors and user mobility at each individual layer of the protocol stack and within each specific problem domain. The preceding approaches have been shown to produce performance gains. In this effort we intend to explore the impact of multi-layer adaptation where several individual adaptation components are composed together, in improving system performance. A fundamental issue is to investigate the relative performance gains that accrue from its use in adaptive handling of resource dynamics, as opposed to single layer adaptation, (barring design constraints in certain environments)? For instance, can the following three adaptation components: adaptive FEC error control, adaptive packet scheduling and channel-aware TCP (e.g. TCP snoop, WTCP) be used to collaboratively improve throughput under different channel conditions?

Other issues include investigating the robustness and sensitivity of multi-layer adaptation as well as its stability and convergence properties. For instance, in an extreme case, in the presence of small deviation from the current network condition, is it possible for the performance gains due to multi-layer adaptation to "disappear"? Similarly, if several adaptation components are composed, would the overall operation converge to the system steady state or may they constantly oscillate and never stabilize in some extreme cases? It is also useful to explore the possibility of design conflicts in adaptation: multiple adaptation schemes have been designed based on different goals, and some of these design goals may be in (partial) conflict. For example, wireless fair scheduling typically treats fairness as the primary design tenet, while many wireless TCP protocols seek to maximize the overall throughput as much as possible. It is well known that ensuring fairness and achieving maximal throughput may be conflicting goals in a generic network topology. The resolution of design conflicts in a scenario where multiple adaptation components are composed poses another design challenge.

Performance and Scalability Prediction

Scalability is a key component of the ubiquitous computing environment of the future. Establishing the scalability of proposed solutions in iMASH presents interesting new challenges. The design of a highly scalable system such as the hospital environment, where a multitude of diverse applications are executing on heterogeneous networks, are notoriously complex to model. They are analytically intractable; abstract simulations of mobile wireless simulations are inaccurate, and detailed simulation models may be computationally impractical for any but small configurations. Further, previous performance studies of mobile and wireless systems have typically emphasized either the application or the network such that only one of the two major system components is modeled in detail and the other component is represented by an abstract model. Thus, networking-oriented simulators (e.g., GloMoSim [13], NS [14], OPNET) will tend to develop a detailed model of the network, possibly including models of the protocol, together with propagation medium and radio models for wireless networks, while representing the application simply as stochastic, or possibly trace-based, traffic streams. On the other hand, application-centric studies tend to develop detailed models of the application, while using an abstract model to represent the network. However, in environments like iMASH where there are closed loop interactions among the application and the network, where changes in network connectivity, for instance, dynamically affect the traffic generated by the application in a statically unpredictable manner and vice versa, such models are inadequate. The ability to incorporate detailed application behavior in a performance prediction environment for mobile, wireless networks, thus appears to have significant value. A central challenge in this regard is the design of multi-paradigm performance prediction technologies, possibly using parallel and hybrid models to predict the performance of dynamic applications in networks with tens of thousands of nodes.

A possible approach to scalable evaluation of applications on very large networks is to use hardware/software-in-the-loop models that interface *physical applications* with a *simulation model* of the network. A challenging requirement of such models is that the simulated components execute sufficiently fast, at least faster than real time [15]; effective support for parallel execution of the simulation models can substantially enhance their applicability in this context. Such a hybrid modeling capability is needed to provide the ability to predict the performance of both the underlying protocols as well as the end applications as a function of network characteristics such as number of nodes, mobility patterns, handoff scenarios etc.

5. Related Work

There are several ongoing projects aimed at supporting nomadic applications. Some not discussed earlier are mentioned briefly here. The Bay Area Research Wireless Access Network (BARWAN) Project [16] seeks to develop a scalable network architecture that can support mobile wireless access across multiple overlay networks while delivering high levels of end-to-end performance to applications. They have developed network protocols for seamless mobility both within a network and across heterogeneous networks, enabled automatic discovery and configuration of local network services.. BARWAN has also developed a general proxy architecture for heterogeneous mobile clients, and built it into a platform for composable infrastructure services, called TACC (transformation, aggregation, caching and customization) [9]. The main goal of TACC is to perform efficient content adaptation for heterogeneous mobile clients. ICEBERG (Internet-based CEllular core BEyond the thiRd Generation) [4, 17] seeks to demonstrate innovative ways to construct a cellular telephone "core" network, replacing traditional mobile switching centers with mobile IP. Monarch [18] seeks to develop networking protocols and protocol interfaces to allow truly seamless wireless and mobile host networking. Its main focus is on mobile IP and ad hoc routing.

NINJA (Ninja Is Not JAva) [19] seeks to develop a distributed middleware architecture to enable programmatic generation and composition of Internet-based services. NINJA's basic approach is based on strongly typed reusable components that are amenable to automated analysis and composition via techniques analogous to those used by compilers. It is in some sense a programming language, except that it is for dynamic creation of wide-area paths rather than executing instructions on a CPU. Mobiware [20] seeks to design, implement and evaluate a mobile middleware toolkit that enables adaptive mobile services. The mobiware toolkit is based on a methodology of open programmable CORBA interfaces and objects representing network resources, for the introduction, control and management of new adaptive services.

6. Conclusion

The need for wireless and mobile access to the network by healthcare providers is but one instance of a universal trend that data access is increasingly becoming wireless and mobile. Collaborative, long distance learning, personalized news delivery, e-commerce are a few instances of significant emerging applications that can similarly exploit the preceding trends. The traditional Internet application model, where each client contacts application servers directly to fetch needed data, does not appear to be optimally suited to support the anytime, anywhere, from any device, access to the Internet that will be necessary to support such applications. In this paper, we propose the use of a middleware infrastructure to: a) help bridge the gap between application servers and nomadic clients as they migrate active sessions among multiple devices; b) balance application requirements and network conditions; and c) enhance the scalability of the network. Significant design challenges that include security and authentication, middleware server discovery, active session handoff, wide area consistency management, and scalability were discussed together with a brief outline of our current approach to addressing these issues as part of the NSF-funded iMASH project

References

1. Bond, A., Gallagher, M., and Indulska, J. *An information model for nomadic environments*. 1998.
2. Bennington, B.J. and Bartel, C.R. *Wireless Andrew: experience building a high speed, campus-wide wireless data network*. in *Proceedings of the Third Annual ACM/IEEE International Conference on Mobile Computing and Networking*. 1997. Budapest, Hungary: ACM, New York, NY, USA.
3. Arnold, K., O'Sullivan, B., Scheifler, R., and Wollrath, A., *The Jini Specification*. 1999: Addison Wesley.
4. Czerwinski, S., Zhao, B., Hodes, T., Joseph, A., and Katz, R. *An Architecture for a Secure Service Discovery Service*. in *MOBICOM '99*. 1999: ACM.
5. Lenoski, D., Laudon, J., Gharachorloo, K., Weber, W.-D., Gupta, A., Hennessy, J., Horowitz, M., and Lam, M.S., *The Stanford DASH Multiprocessor*. IEEE Computer, 1992. 25(3): p. 63--79.
6. Satyanarayanan, M., Kistler, J.J., Kumar, P., Okasaki, M.E., Siegel, E.H., and Steere, D.C., *Coda: a highly available file system for a distributed workstation environment*. IEEE Transactions on Computers, 1990. 39(4): p. 447-59.
7. Noble, B.D., Satyanarayanan, M., Narayanan, D., Tilton, J.E., Flinn, J., and Walker, K.R. *Agile application-aware adaptation for mobility*. in *16th ACM Symposium on Operating Systems Principles*. 1997. Saint Malo, France.
8. Joseph, A., Tauber, J., and Kaashock, M.F. *Building Reliable Mobile-Aware Applications Using the Rover Toolkit*. in *Second ACM International Conference on Mobile Computing and Networking (MoBiCom '96)*. 1996.
9. Fox, A., Gribble, S.D., Chawathe, Y., and Brewer, E.A., *Adapting to network and client variation using infrastructural proxies: lessons and perspectives*. IEEE Personal Communications, 1998. 5(4): p. 10-19.
10. Feldmeier, D.C., McAuley, A.J., Smith, J.M., Bakin, D.S., Marcus, W.S., and Raleigh, T.M., *Protocol boosters*. IEEE Journal on Selected Areas in Communications, 1998. 16(3): p. 437-44.
11. Sudame, P. and Badrinath, B.R. *Transformer tunnels: a framework for providing route-specific adaptations*. in *USENIX 1998 Annual Technical Conference*. 1998. New Orleans, LA, USA.
12. Harreld, M., Valentino, D.J., Liu, B.J., El-Saden, S.M., and Duckwiler, G.R., *Advanced Workstation For Diagnostic Neuroradiology: The Neuro Navigator*. Radiology, 1997. 205: p. 743.
13. Zeng, X., Bagrodia, R., and Gerla, M. *GloMoSim: a library for parallel simulation of large-scale wireless networks*. in *Twelfth Workshop on Parallel and Distributed Simulation PADS '98*. 1998. Banff, Alta., Canada.
14. *Wireless and Mobility Extensions to NS*, , http://www.monarch.cs.cmu.edu/cmu-ns.html.
15. reference deleted for double blind
16. Brewer, E.A., Katz, R.H., Chawathe, Y., Gribble, S.D., Hodes, T., Giao, N., Stemm, M., Henderson, T., Amir, E., Balakrishnan, H., Fox, A., Padmanabhan, V.N., and Seshan, S., *A network architecture for heterogeneous mobile computing*. IEEE Personal Communications, 1998. 5(5): p. 8-24.
17. Ludwig, R., Konrad, A., and Joseph, A. *Optimizing the End-to-End Performance of Reliable Flows over Wireless Links*. in *MOBICOM '99*. 1999: ACM.

18. Johnson, D.B. and Maltz, D.A., *Truly seamless wireless and mobile host networking. Protocols for adaptive wireless and mobile networking*. IEEE Personal Communications, 1996. 3(1): p. 34-42.
19. Gribble, S.D., Welsh, M., Brewer, E.A., and Culler, D. *The MultiSpace: an Evolutionary Platform for Infrastructural Services*. in *Usenix '99*. 1999.
20. Angin, O., Campbell, A.T., Kounavis, M.E., and Liao, R.R.F., *The mobiware toolkit: programmable support for adaptive mobile networking*. IEEE Personal Communications, 1998. 5(4): p. 32-43.