

The Essence of Pervasive Computing in the Environment Integrated with Human Users

Venkata Rao Barige
Assistant Professor
Department of MCA

V. R. Siddhartha Engg College, Vijayawada, A.P, India

B.D.C.N Prasad
Professor & Head
Department of MCA
PVPSIT, Vijayawada, A.P, India

Abstract

This paper realizes the state of affairs in computer systems research positioned by the emerging field of pervasive computing. It first analyse the relationship of this new field to its predecessors: distributed systems and mobile computing. It then identifies four new research thrusts: impressive use of smart spaces, invisibility, localized scalability, and masking uneven conditioning. Next, it summarizes a couple of hypothetical pervasive computing scenarios, and uses them to identify key capabilities missing from today's systems. The paper appressed with a discussion of the research necessary to develop these capabilities.

Keywords: Pervasive Computing, Smart devices, Research thrusts, key capabilities, Relationship, Invisibility, Scalability

I. INTRODUCTION

“The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it.” So began Mark Weiser’s seminal 1991 paper [44] that described his vision of ubiquitous computing, now also called pervasive computing. The essence of that vision was the creation of environments saturated with computing and communication capability, yet gracefully integrated with human users. When articulated, this was a vision too far ahead of its time — the hardware technology needed to achieve it simply did not exist. Not surprisingly, the implementation attempted by Weiser and his colleagues at Xerox PARC fell short. After a decade of hardware progress, many critical elements of pervasive computing that were exotic in 1991 are now viable commercial products: handheld and wearable computers; wireless LANs; and devices to sense and control appliances. We are now better positioned to begin the quest for Weiser’s vision. Pervasive computing projects have emerged at major universities and in industry. Examples at universities include Project Aura at Carnegie Mellon, Endeavour at UC Berkeley, Oxygen at MIT, and Portalano at Washington. Industry examples include work at AT&T Research in Cambridge, U.K. and at the IBM TJ Watson Research Center. Each of these projects addresses a different mix of issues in pervasive computing, and a different blend of nearterm and far-term goals. Together, they represent a broad communal effort to make pervasive computing a reality. The goal of this paper is to help us understand the challenges in computer systems research posed by pervasive computing. We begin by examining its relationship to the closely-related fields of distributed systems and mobile computing. Next, we sketch two pervasive computing scenarios, and ask why they are fiction rather than fact today. From that starting point, we delve deeper into some key research problems. To preserve focus on computer systems issues, we avoid digressions into other areas important to pervasive computing such as human-computer interaction, expert systems and software agents

II. RELATED FIELDS

Pervasive computing represents a major evolutionary step in a line of work dating back to the mid-1970’s. Two distinct earlier steps in this evolution are distributed systems and mobile computing. Some of the technical problems in pervasive computing correspond to problems already identified and studied earlier in the evolution. In some of those cases, existing solutions apply directly; in other cases, the demands of pervasive computing are sufficiently different that new solutions have to be sought. There are also new problems introduced by pervasive computing that have no obvious mapping to problems studied earlier. In the rest of this section, we try to sort out this complex intellectual relationship and to develop a taxonomy of issues characterizing each phase of the evolution

A. Distributed Systems:

The field of distributed systems arose at the intersection of personal computers and local area networks. The research that followed from the mid-1970’s through the early 1990’s created a conceptual framework and algorithmic base that has proven to be of enduring value in all work involving two or more computers connected by a network — whether mobile or static, wired or wireless, sparse or pervasive. This body of knowledge spans many areas that are foundational to pervasive computing and is now well codified in textbooks [8, 19, 20]: • remote communication, including protocol layering, remote procedure call [3], the use of timeouts, and the use of end-to-end arguments in placement of functionality [28]. • fault tolerance, including atomic transactions, distributed and nested transactions, and two-phase commit [13]. • high availability, including optimistic and pessimistic replica control [9], mirrored execution [4], and optimistic recovery [37]. • remote information access, including caching, function

shipping, distributed file systems, and distributed databases [30]. • security, including encryption-based mutual authentication and privacy [23].

B. Mobile Computing:

The appearance of full-function laptop computers and wireless LANs in the early 1990s led researchers to confront the problems that arise in building a distributed system with mobile clients. The field of mobile computing was thus born. Although many basic principles of distributed system design continued to apply, four key constraints of mobility forced the development of specialized techniques. These constraints are: unpredictable variation in network quality, lowered trust and robustness of mobile elements, limitations on local resources imposed by weight and size constraints, and concern for battery power consumption [31]. Mobile computing is still a very active and evolving field of research, whose body of knowledge awaits codification in textbooks. The results achieved so far can be grouped into the following broad areas: • mobile networking, including Mobile IP [2], ad hoc protocols [27], and techniques for improving TCP performance in wireless networks [1, 5]. • mobile information access, including disconnected operation [17], bandwidth-adaptive file access [21], and selective control of data consistency [38, 39]. • support for adaptative applications, including transcoding by proxies [12] and adaptive resource management [24]. • system-level energy saving techniques, such as energyaware adaptation [11], variable-speed processor scheduling [45], and energy-sensitive memory management [18]. • location sensitivity, including location sensing [42, 43] and location-aware system behavior [32, 35, 41].

C. Pervasive Computing:

Earlier in this paper, we characterized a pervasive computing environment as one saturated with computing and communication capability, yet so gracefully integrated with users that it becomes a “technology that disappears.” Since motion is an integral part of everyday life, such a technology must support mobility; otherwise, a user will be acutely aware of the technology by its absence when he moves. Hence, the research agenda of pervasive computing subsumes that of mobile computing, but goes much further. Specifically, pervasive computing incorporates four additional research thrusts into its agenda as shown Fig 1.

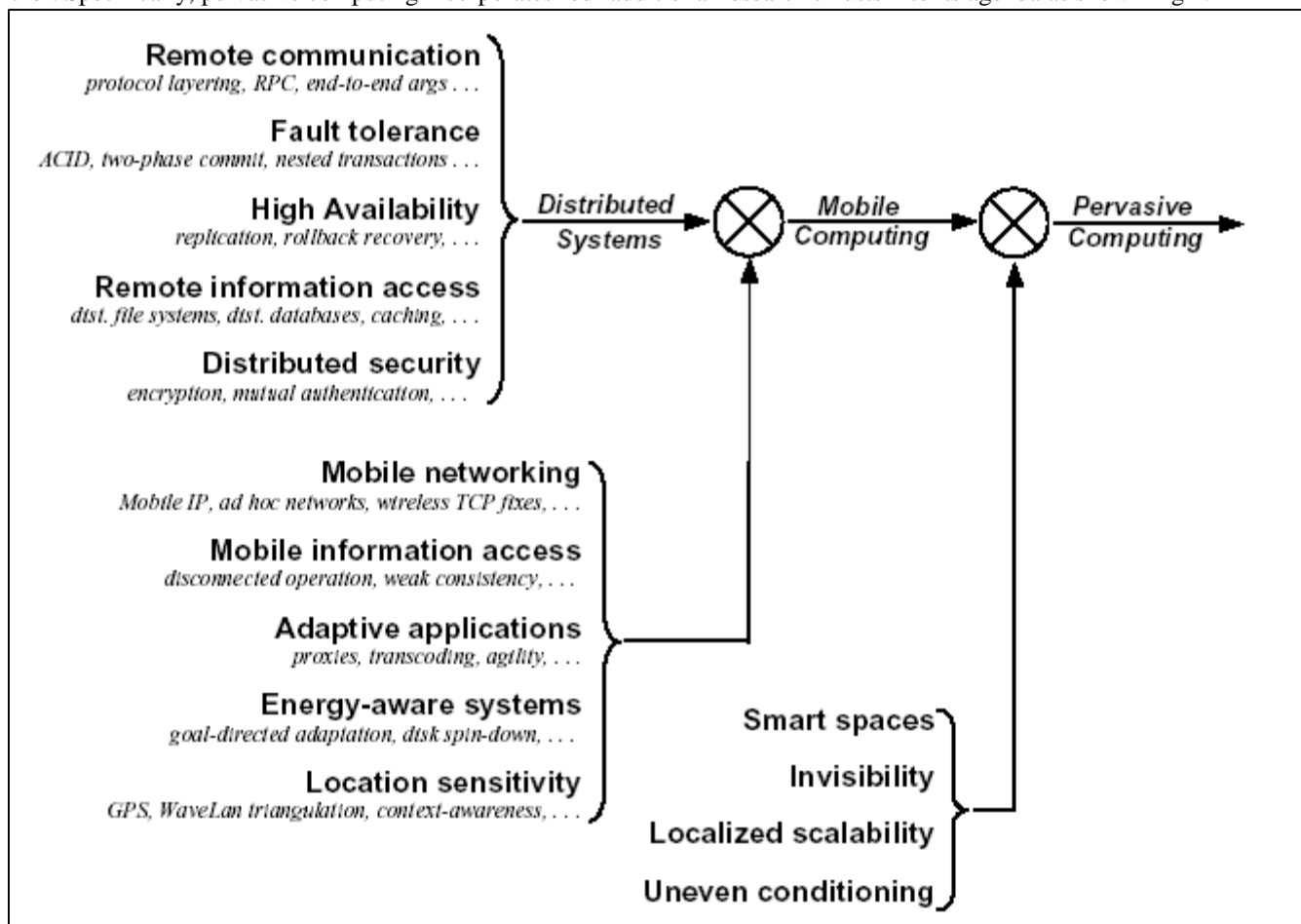


Fig. 1: Taxonomy of Computer Systems Research Problems in Pervasive Computing

1) Effective Use of Smart Spaces:

The first research thrust is the effective use of smart spaces. A space may be an enclosed area such as a meeting room or corridor, or it may be a well-defined open area such as a courtyard or a quadrangle. By embedding computing infrastructure in building infrastructure, a smart space

brings together two worlds that have been disjoint until now [16]. The fusion of these worlds enables sensing and control of one world by the other. A simple example of this is the automatic adjustment of heating, cooling and lighting levels in a room based on an occupant's electronic profile. Influence in the other direction is also possible — software on a user's computer may behave differently depending on where the user is currently located. Smartness may also extend to individual objects, whether located in a smart space or not.

2) *Invisibility:*

The second thrust is invisibility. The ideal expressed by Weiser is complete disappearance of pervasive computing technology from a user's consciousness. In practice, a reasonable approximation to this ideal is minimal user distraction. If a pervasive computing environment continuously meets user expectations and rarely presents him with surprises, it allows him to interact almost at a subconscious level [46]. At the same time, a modicum of anticipation may be essential to avoiding a large unpleasant surprise later — much as pain alerts a person to a potentially serious future problem in a normally-unnoticed body part.

3) *Localized Scalability:*

The third research thrust is localized scalability. As smart spaces grow in sophistication, the intensity of interactions between a user's personal computing space and his surroundings increases. This has severe bandwidth, energy and distraction implications for a wireless mobile user. The presence of multiple users will further complicate this problem. Scalability, in the broadest sense, is thus a critical problem in pervasive computing. Previous work on scalability has typically ignored physical distance — a web server or file server should handle as many clients as possible, regardless of whether they are located next door or across the country. The situation is very different in pervasive computing. Here, the density of interactions has to fall off as one moves away — otherwise both the user and his computing system will be overwhelmed by distant interactions that are of little relevance. Although a mobile user far from home will still generate some distant interactions with sites relevant to him, the preponderance of his interactions will be local. Like the inverse square laws of nature, good system design has to achieve scalability by severely reducing interactions between distant entities. This directly contradicts the current ethos of the Internet, which many believe heralds the "death of distance."

4) *Masking Uneven Conditioning:*

The fourth thrust is the development of techniques for masking uneven conditioning of environments. The rate of penetration of pervasive computing technology into the infrastructure will vary considerably depending on many non-technical factors such as organizational structure, economics and business models. Uniform penetration, if it is ever achieved, is many years or decades away. In the interim, there will persist huge differences in the "smartness" of different environments — what is available in a well-equipped conference room, office, or classroom may be more sophisticated than in other locations. This large dynamic range of "smartness" can be jarring to a user, detracting from the goal of making pervasive computing technology invisible. One way to reduce the amount of variation seen by a user is to have his personal computing space compensate for "dumb" environments. As a trivial example, a system that is capable of disconnected operation is able to mask the absence of wireless coverage in its environment. Complete invisibility may be impossible, but reduced variability is well within our reach.

III. CONCLUSIONS

Pervasive computing will be a fertile source of challenging research problems in computer systems for many years to come. Solving these problems will require us to broaden our discourse on some topics, and to revisit long-standing design assumptions in others. We will also have to address research challenges in areas outside computer systems. These areas include human-computer interaction (especially multi-modal interactions and human-centric hardware designs), software agents (with specific relevance to high-level proactive behavior), and expert systems and artificial intelligence (particularly in the areas of decision making and planning). Capabilities from these areas will need to be integrated with the kinds of computer systems capabilities discussed in this paper. Pervasive computing will thus be the crucible in which many disjoint areas of research are fused. When describing his vision, Weiser was fully aware that attaining it would require tremendous creativity and effort by many people, sustained over many years. The early decades of the 21st century will be a period of excitement and ferment, as new hardware technologies converge with research progress on the many fundamental problems discussed in this paper. Like the Frontier of the American West in the early 19th century, pervasive computing offers new beginnings for the adventurous and the restless — a rich open space where the rules have yet to be written and the borders yet to be drawn.

ACKNOWLEDGEMENTS

The ideas in this paper were formed over an extended period of time and benefited from the input of many individuals. In particular, I would like to thank my colleagues on the Aura team (David Garlan, Raj Reddy, Dan Siewiorek, Asim Smailagic and Peter Steenkiste) for their many valuable thoughts, suggestions and insights. I would also like to acknowledge the members of the Coda and Odyssey projects for the many insights that I have gained in working with them. Discussions with Dave Clark and Mike Dertouzos of MIT were helpful in formulating the concept of localized scalability, as discussed in Section 2.3. I would like to thank Sandeep Gupta, guest editor of this special issue, for giving me the opportunity to express my thoughts on pervasive computing. I would also like to thank a number of people who read early versions of this paper and provided valuable feedback on its content and presentation: Mary Baker, Rajesh Balan, Maria Ebling, Michalis Faloutsos, Jason Flinn, David Garlan, Ira Greenberg, Guerny Hunt, Hui Lei, Alan Messer, Dejan Milojovic, Dushyanth Narayanan, and SoYoung Park. Any remaining errors or omissions are, of course, entirely my responsibility. This research was supported by the National Science Foundation (NSF) under contracts CCR-9901696 and ANI-0081396, the Defense Advanced Projects Research Agency (DARPA) and the U.S. Navy (USN) under contract N660019928918, IBM Corporation, Compaq Corporation and Nokia Corporation. The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of the NSF, DARPA, USN, IBM, Compaq, Nokia, or the U.S. government.

REFERENCES

- [1] Bakre, A., Badrinath, B.R. Handoff and System Support for Indirect TCP/IP. In Proceedings of the Second Usenix Symposium on Mobile & Location-Independent Computing. Ann Arbor, MI, April, 1995.
- [2] Bhagwat, P., Perkins, C., Tripathi, S. Network Layer Mobility: An Architecture and Survey. IEEE Personal Communications 3(3), June, 1996.
- [3] Birrell, A.D., Nelson, B.J. Implementing Remote Procedure Calls. ACM Transactions on Computer Systems 2(1), February, 1984.

- [4] Borg, A., Blau, W., Graetsch, W. Fault Tolerance Under Unix. *ACM Transactions on Computer Systems* 7(1), February, 1989.
- [5] Brewer, E.A., Katz, R.H., Chawathe, Y., Gribble, S.D., Hodes, T., Nguyen, G., Stemm, M., Henderson, T., Amir, E., Balakrishnan, H., Fox, A., Padmanabhan, V.N., Seshan, S. A Network Architecture for Heterogeneous Mobile Computing. *IEEE Personal Communications* 5(5), October, 1998.
- [6] Brodersen, R.W. InfoPad — past, present and future. *Mobile Computing and Communications Review* 3(1), January, 1999.
- [7] Case, J., Fedor, M., Schoffstall, M., Davin, J. A Simple Network Management Protocol. Internet Engineering Task Force (RFC 1157), 1990.
- [8] Couloris, G., Dollimore, J., Kindberg, T. *Distributed Systems Concepts and Design* (Third Edition). Addison-Wesley, 2001.
- [9] Davidson, S.B., Garcia-Molina, H., Skeen, D. Consistency in Partitioned Networks. *ACM Computing Surveys* 17(3), September, 1985.
- [10] Ellis, C.S. The Case for Higher-Level Power Management. In *The 7th IEEE Workshop on Hot Topics in Operating Systems*. Rio Rico, AZ, March, 1999.
- [11] Flinn, J., Satyanarayanan, M. Energy-aware Adaptation for Mobile Applications. In *Proceedings of the 17th ACM Symposium on Operating Systems and Principles*. Kiawah Island, SC, December, 1999.
- [12] Fox, A., Gribble, S.D., Brewer, E.A., Amir, E. Adapting to Network and Client Variability via On-Demand Dynamic Distillation. In *Proceedings of the Seventh International ACM Conference on Architectural Support for Programming Languages and Operating Systems*. Cambridge, MA, October, 1996.
- [13] Gray, J., Reuter, A. *Transaction Processing: Concepts and Techniques*. Morgan Kaufman, 1993.
- [14] Itoi, N., Honeyman, P. Practical Security Systems with Smartcards. In *The 7th IEEE Workshop on Hot Topics in Operating Systems*. Rio Rico, AZ, March, 1999.
- [15] Jain, A., Hong, L., Pankanti, S. Biometric Identification. *Communications of the ACM* 43(2), February, 2000.
- [16] Katz, R.H., Long, D., Satyanarayanan, M., Tripathi, S. Workspaces in the Information Age. In *Report of the NSF Workshop on Workspaces in the Information Age*. Leesburg, VA, October, 1996. Available at <http://www.cs.berkeley.edu/~randy/NSFWS>.
- [17] Kistler, J.J., Satyanarayanan, M. Disconnected Operation in the Coda File System. *ACM Transactions on Computer Systems* 10(1), February, 1992.
- [18] Lebeck, A.R., Fan, X., Zheng, H., Ellis, C.S. Power Aware Page Allocation. In *Proceedings of the Ninth International Conference on Architectural Support for Programming Languages and Operating Systems*. November, 2000.
- [19] Lynch, N.A. *Distributed Algorithms*. Morgan Kaufmann, 1993.
- [20] Mullender, S.J. (editor). *Distributed Systems*. Addison-Wesley, 1993.
- [21] Mummert, L.B., Ebling, M.R., Satyanarayanan, M. Exploiting Weak Connectivity for Mobile File Access. In *Proceedings of the 15th ACM Symposium on Operating Systems Principles*. Copper Mountain Resort, CO, December, 1995.
- [22] Nahrsted, K., Chu, H., Narayan, S. QoS-aware Resource Management for Distributed Multimedia Applications. *Journal on High-Speed Networking* 7(3/4), 1998.
- [23] Needham, R.M and Schroeder, M.D. Using Encryption for Authentication in Large Networks of Computers. *Communications of the ACM* 21(12), December, 1978.
- [24] Noble, B.D., Satyanarayanan, M., Narayanan, D., Tilton, J.E., Flinn, J., Walker, K.R. Agile Application-Aware Adaptation for Mobility. In *Proceedings of the 16th ACM Symposium on Operating Systems Principles*. Saint-Malo, France, October, 1997.
- [25] Energy-Efficient Technologies for the Dismounted Soldier Board on Army Science and Technology, National Research Council, Washington, DC, 1997.
- [26] Parnas, D.L. On the Criteria to be Used in Decomposing Systems into Modules. *Communications of the ACM* 15(12), December, 1972.
- [27] Royer, E.M., Toh, C.K. A Review of Current Routing Protocols for Ad Hoc Mobile Wireless Networks. *IEEE Personal Communications* 6(2), April, 1999.
- [28] Saltzer, J.H., Reed, D.P., Clark, D.D. End-to-End Arguments in System Design. *ACM Transactions on Computer Systems* 2(4), November, 1984.
- [29] Satyanarayanan, M. Caching Trust Rather Than Content. *Operating System Review* 34(4), October, 2000.
- [30] Satyanarayanan, M. A Survey of Distributed File Systems. In Traub, J.F., Grosz, B., Lampson, B., Nilsson, N.J. (editors), *Annual Review of Computer Science*. Annual Reviews, Inc, 1989.
- [31] Satyanarayanan, M. Fundamental Challenges in Mobile Computing. In *Proceedings of the Fifteenth ACM Symposium on Principles of Distributed Computing*. Philadelphia, PA, May, 1996.
- [32] Schilit, B., Adams, N., Want, R. Context-Aware Computing Applications. In *Proceedings of the Workshop on Mobile Computing Systems and Applications*. Santa Cruz, CA, December, 1994.
- [33] Schmidt, B.K., Lam, M.S., Northcutt, J.D. The Interactive Performance of SLIM: a Stateless, Thin-Client Architecture. In *Proceedings of the 17th ACM Symposium on Operating Systems and Principles*. Kiawah Island, SC, December, 1999.
- [34] Smailagic, A., Siewiorek, D.P. Modalities of Interaction with CMU Wearable Computers. *IEEE Personal Communications* 3(1), February, 1996.
- [35] Spreitzer, M., Theimer, M. Providing Location Information in a Ubiquitous Computing Environment. In *Proceedings of the 14th ACM Symposium on Operating System Principles*. December, 1993.
- [36] Steiner, J.G., Neuman, G., Schiller, J.I. Kerberos: An Authentication Service for Open Network Systems. In *Proceedings of the Winter 1988 USENIX Technical Conference*. Dallas, TX, February, 1988.
- [37] Strom, R.E., Yemini, S. Optimistic Recovery in Distributed Systems. *ACM Transactions on Computer Systems* 3(3), August, 1985.
- [38] Tait, C.D., Duchamp, D. An Efficient Variable-Consistency Replicated File Service. In *Proceedings of the USENIX File Systems Workshop*. Ann Arbor, MI, May, 1992.
- [39] Terry, D.B., Theimer, M.M., Petersen, K., Demers, A.J., Spreitzer, M.J., Hauser, C.H. Managing Update Conflicts in a Weakly Connected Replicated Storage System. In *Proceedings of the 15th ACM Symposium on Operating Systems Principles*. Copper Mountain Resort, CO, December, 1995.
- [40] Truman, T.E., Pering, T., Doering, R., Brodersen, R.W. The InfoPad Multimedia Terminal: A Portable Device for Wireless Information Access. *IEEE Transactions on Computers* 47(10), October, 1998.
- [41] Voelker, G.M., Bershady, B.N. Mobisaic: An Information System for a Mobile Wireless Computing Environment. In *Proceedings of the Workshop on Mobile Computing Systems and Applications*. Santa Cruz, CA, December, 1994.
- [42] Want, R., Hopper, A., Falcao, V., Gibbons, J. The Active Badge Location System. *ACM Transactions on Information Systems* 10(1), January, 1992.
- [43] Ward, A., Jones, A., Hopper, A. A New Location Technique for the Active Office. *IEEE Personal Communications* 4(5), October, 1997.
- [44] Weiser, M. The Computer for the 21st Century. *Scientific American*, September, 1991.
- [45] Weiser, M., Welch, B., Demers, A., Shenker, S. Scheduling for reduced CPU energy. In *Proceedings of the First USENIX Symposium on Operating System Design and Implementation*. Monterey, CA, November, 1994.
- [46] Weiser, M, Brown, J.S. The Coming Age of Calm Technology. In Denning, P.J., Metcalfe, R.M. (editors), *Beyond Calculation: The Next Fifty Years of Computing*. Copernicus, 1998.
- [47] M.Satyanarayanan, *Pervasive Computing: Vision and Challenges*, IEEE Personal communications, 2001