

Cross-Layer Architectures for Autonomic Communications

Mohammad Abdur Razzaque,^{1,2} Simon Dobson,¹ and Paddy Nixon¹

Published online: 3 January 2007

Layered architectures are not flexible enough to cope with the dynamics of wireless dominated next generation communications. Cross-layer architectures may provide a more flexible solution; breaks the traditional structure by allowing interactions between two or more non-adjacent layers. This paper review the cross-layer approach to network architecture and compare the different cross-layering architectures, observing that most current approaches depend purely on local information and provide only poor and inaccurate information gathering at the global scale. This paper also explores the possible use of cross-layering architectures in autonomic communications and the potential importance of new cross-layer architectures with a hybrid local and global view for autonomic communications.

KEY WORDS: Layered architectures; cross-layering; self-* properties.

1. INTRODUCTION

The traditional layered design of network protocols is insufficiently flexible to cope with the dynamics of wireless-dominated next-generation communications. Recent studies [1] show that careful exploitation of some protocol interactions that cross the normal layer boundaries can lead to more efficient performance of the transmission stack—and hence to better application-layer performance—in a number of different wireless scenarios.

Cross-layer design breaks away from traditional network design, where each layer of the protocol stack operates independently. In the cross-layer approach information is exchanged between different layers of the protocol stack, and end-to-end performance is optimized by adapting each layer against this information.

¹Systems Research Group, School of Computer Science and Informatics UCD, Dublin, Ireland.

² To whom correspondence should be addressed at Systems Research Group, School of Computer Science and Informatics UCD, Dublin, Ireland. E-mail: abdur.razzaque@ucd.ie.

Cross-layering is not the simple replacement of a layered architecture, nor is it the simple combination of layered functionality; instead it breaks the boundaries between information abstractions to improve end-to-end transportation.

It is clear from the recent initiatives in autonomic computing and autonomic communications [2, 3] that there is a need to make future networks *self-behaving*, in the sense that they work in an optimal way with “endogenous” management and control, with minimum human perception and intervention. To attain such a self-behaving system within existing strictly-layered approaches may be possible, but will not (we claim) leverage all the possible optimisations. We consider cross-layer architectures to be better suited to achieving the self-optimisation, self-configuration and the other “self-*” properties targeted by autonomic approaches.

A number of proposals for cross layer designs and their corresponding architectures have been published in the literature. Alongside the categorizations mentioned in [1], all of the existing cross-layer design architectures could be classified according to how they are getting the information for the optimisations: (i) architectures based on local information (from the node and its different layers); and (ii) architectures based on local and global information (from the node, its different layers and from neighbors).

Most existing architectures (including GRACE(Global Resource Adaptation through Co-opEration) [4], WIDENS (Wireless DEployable Network System) [5], MobileMan [6]) are based on a local view of the state of, and constraints on, the network: only the CrossTalk [7] is based on global view (even partially). On the other hand, POEM (Performance-Oriented Model) [8] is the only architecture considering self-optimisation that could be helpful for autonomic communications.

The main objective of this paper is to provide a review of cross-layering approaches in next-generation communications, and their differences from existing architectures. A secondary objective is to explore the possible use of cross-layering architectures in autonomic communications and the potential importance of new cross-layer architectures with a hybrid local and global view for autonomic communications.

The rest of the paper is structured as follows. Section 2 describes cross-layered architectures in the abstract, which is then used to inform a survey of existing approaches in section 3. Section 4 draws-out some of the issues in autonomic systems and discusses the possible use of cross-layer architecture in autonomic communications. Section 5 concludes with some directions for future work.

2. RATIONALE FOR CROSS-LAYER NETWORKING

Layering is the dominant design methodology in communications protocol stacks, but this dominance of strict layering is being threatened by next-generation wireless-dominated networking.

2.1. What is Cross-Layer Networking?

Cross-layer design breaks away from the traditional network design, where each layer of the protocol stack operates independently and information is exchanged only between adjacent layers via narrow interfaces. Information in cross-layer architecture is exchanged between non-adjacent layers of the protocol stack, typically using a broader and more open data format, and end-to-end performance is optimised by adapting to this information at each protocol layer. Cross-layering is therefore not the simple replacement of a layered architecture, nor is it the simple combination of layered functionality: instead, cross-layering attempts to share information amongst different layers, which can be used as input for algorithms, for decision processes, and adaptations.

Existing cross-layer interaction largely focuses on direct interactions between the protocols by involving only two or three layers and introducing shortcuts between protocols [9, 10], and most focus simply on the energy constraint and (to a lesser extent) certain forms of security [6].

2.2. Why Cross-Layering?

Cross-layer design can therefore play an important role for the next-generation wireless systems, featured by all IP-based protocol stack, heterogeneous access networks, and multimedia data traffic. We can look at the motivation for cross layering in communications in two ways, from a general communications viewpoint and then from a more targeted wireless viewpoint.

2.2.1. General Communication Viewpoint

One obvious shortcoming of the two classical network reference models, OSI and TCP/IP, is the lack of information sharing between the protocol layers. This hampers optimal performance of the networks, since shared layer information is the prerequisite for many forms of performance optimisation. Cross-layer systems shift the research landscape away from optimizing the performance of individual layers, and instead treat optimisation as a problem for the entire stack. The technique consists of taking into account information available from different levels [10], not necessarily adjacent, in order to create a system much more sensitive to its environment, and load.

OSI (Open System Interconnection) and TCP/IP support a bottom-up approach driven by physical and network constraints, which makes it hard to capture and respond to (top-down) user demands or requirements. Cross-layer design can help to capture these concerns by providing a more uniform framework within which to capture and disseminate concerns at different semantic levels.

Introducing a single co-located layer for various adaptation tasks would be too complex and heavyweight, and even then would be inadequate: QoS(Quality of

Service) adaptation requires all layers' participation [11]. A co-operative solution involving co-ordinating the individual adaptations of multiple layers would lead to a more flexible approach, although introducing the potential for feature interaction and instability.

In addition to performance improvements at lower layers, cross-layering allows us to design new kinds of applications. Specially affected are distributed applications and applications sensitive to changing network conditions such as QoS-sensitive multimedia applications.

2.2.2. *Wireless Networking Viewpoint*

The assumptions in the wired IP stack are inadequate for wireless networking, and TCP is known to suffer from performance degradation in mobile wireless environments. This is because such environments are prone to packet losses due to high bit error rates and mobility-induced disconnections. TCP interprets packet losses as an indication of congestion and (inappropriately) invokes congestion control mechanisms, which leads to degraded performance. With the help of cross-layering this problem could be solved [9].

The combination of scarce radio resource and limited power necessitate the optimisation of network performance, but such optimisation can hardly be met in the sub-optimal wired architecture with strict layering. Using cross-layering better optimisation is possible as shown in [12].

In traditional networks, the Link Layer is for point-to-point communications, while the Transport Layer is for end-to-end communications across various links. In short-range networks these two concerns collapse into one: peer-to-peer communications mostly take place in the point-to-point level. Using cross-layer design, duplicate efforts from each related layer can be avoided [13].

Wireless networks offer several possibilities for opportunistic communication that cannot be exploited sufficiently in a strictly layered design. Furthermore, the wireless medium offers some new modalities of communication the layered architectures do not accommodate, for example making the physical layer capable of receiving multiple packets at the same time [11].

3. EXISTING CROSS-LAYER ARCHITECTURES

Research on cross-layer networking is still at a very early stage, and no consensus exists on a generic cross layer infrastructure or architecture. However, the importance of a good and sound architecture to handle the proliferation of cross-layer operations in wireless as well other communications media is clear, especially in autonomic systems for which properties need to be specified and maintained with minimal manual configuration and intervention [14]. A number of proposals for cross layer designs and their corresponding architectures have been published in the literature. Most of these proposals are based on one of the

basic categories mentioned in [1]. In this paper we are interested in looking at cross-layering architectures in terms of how and from where they gather different cross-layer and optimisation related information. The possible candidates are: (i) Architectures based on local information and (ii) Architectures based on local and global information (from a single node, its different layers, and the states of its neighbors).

3.1. Architectures with Local View of the Network

Most of the existing cross-layer architectures depend on the node-based local view or state information of the network or system in decision making.

The “Interlayer Signalling Pipe” (Fig. 1) [15] stores cross-layer information in the Wireless Extension Headers of IPv6 packets. This header-based method makes use of IP data packets as in-band message carriers with no need to use a dedicated messaging protocol. Normally an IP data packet can only be processed layer-by-layer, however, and this can lead to inefficiency: layer-by-layer signaling constructs a bottom-to-top pipe, which seems excessive in most cases.

In the Internet Control Message Protocol (ICMP) [16] message based architecture, it propagates information across different layers. Since a message could be generated from any layer and then terminates at a higher layer, cross-layer signaling is carried out through these selected “holes,” rather than the “pipe” in [15], as shown in Fig. 2. This appears to be more flexible and efficient than the Interlayer Signaling Pipe. However, ICMP messages encapsulated by IP packets have to pass by the network layer even if the signaling is only desired between the link layer and the application layer.

In the local profile based architecture shown in Fig. 3, cross-layer information is abstracted from each related layer respectively and stored in separate profiles within a Mobile Host (MH) record [13]. Other interested layers can then select

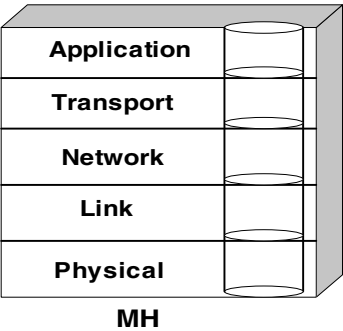


Fig. 1. Pipe architecture [15].

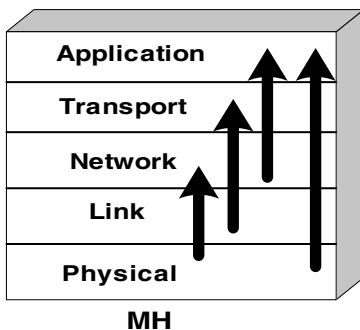


Fig. 2. ICMP based architecture [16].

profiles to fetch any desired information. This is flexible since profile formats can be tailored to specific layers, which in turn can access to the information directly.

In the architecture of [17], channel and link information from the physical and link layers are gathered, abstracted and managed by a third party, the distributed Wireless Channel Information (WCI) servers. Interested applications access the WCI servers to retrieve their desired information as shown in Fig. 4. As a network service, it is complementary to the former schemes within an MH although some overheads would be incurred on the air and interfaces have to be defined between the MH, WCI server and application servers.

In [18] a multi-layer architecture (Fig. 4) for advanced mobility support based on cross-layering is presented. Possible support includes co-ordinated mobility management of different levels for various mobility types, fast/seamless handoffs, and QoS adaptation to the mobility-incurred context changes such as heterogeneous networks. Active cross-layer interactions play a crucial role to enable and facilitate such extended functions and improved performance.

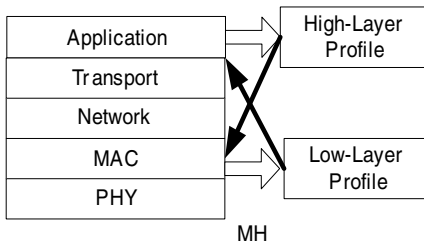


Fig. 3. Local profile-based architecture [17].

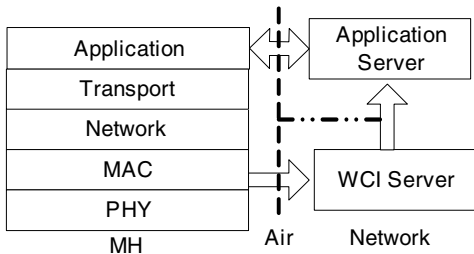


Fig. 4. Service-based architecture [18].

GRACE [4] is a cross layer adaptation framework (Fig. 5). All system components (hardware, network, and operating system) and applications are allowed to be adaptive. These adaptive entities co-operate with each other to achieve a system-wide optimal configuration, for example to maximize system utility, in the presence of changes in the available resources or application demands. However, its cross-layer approach includes no explicit consideration of cross layering within the networking layers or protocol stack. WIDENS [5] has been proposed (Fig. 6) with an aim to acquire the interoperability, cross layering and re-configurability at the same time. This cross layering architecture seems a promising one where protocol optimisation is based on the local state information but it is still in the validation stage and so lacks any real measurement of efficiency especially in terms of performance.

The MobileMan [6] architecture presents (Fig. 7), along with the strict layering, a core component, *Network Status*, which functions as a repository for information that uniformly manages the cross-layer interaction while respecting

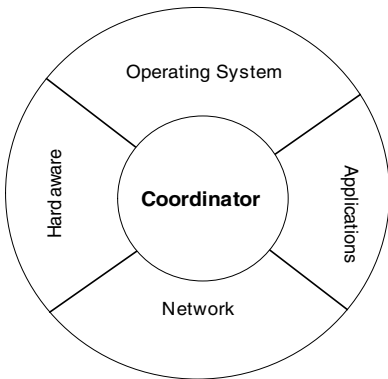


Fig. 5. GRACE’s hierarchical adaptation [4].

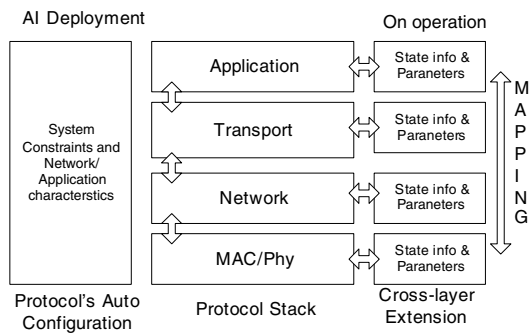


Fig. 6. WIDENS architecture [5].

the principle of dividing functionalities and responsibilities in layers. The approach aims to optimize overall network performance with respect to local state information by increasing local interaction among protocols, decreasing remote communications, and consequently saving network bandwidth. Performance improvement verifications are yet to be published. ECLAIR [19] is a local-view-based, efficient cross-layer architecture (Fig. 8) for wireless protocol stacks. Along with legacy protocol stack it consists of two main components: an *Optimizing Sub-System*,

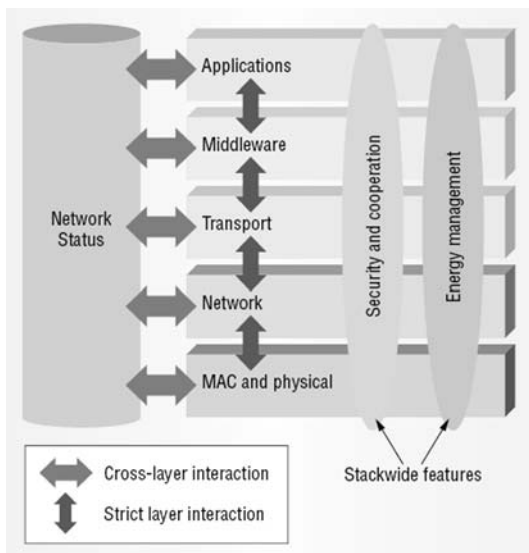


Fig. 7. Mobile man architecture [6].

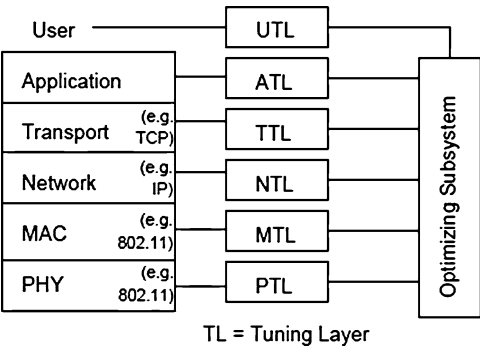


Fig. 8. ECLAIR architecture [19].

the cross layer engine which contains many *Protocol Optimizers*, which are the “intelligent” components of it, and *Tuning Layers* provide the necessary APIs to the protocol optimizers for interacting with various layers and manipulating the protocol data structures. There is no processing overhead on the existing stack since the optimizing subsystem executes in parallel to the protocol stack.

POEM [8] is perhaps the first initiative towards developing a cross-layer based self-optimizing protocol stack specifically for autonomic communication. For the optimisation purposes it utilises local state information. The basic design criterion is self-optimisation is a control-plane issue where the normal functions of the protocol stack should not be compromised, with added cross-layer benefits being layered on top. The system is being investigated both formally and through simulation.

CATNIP (Context-Aware Transport/Network Internet Protocol) [20] uses cross-layering approach to improving features such as packet loss and retrieval time. It does not provide any cross-layering architecture; instead, it is an integrated protocol, which uses application-layer knowledge (for example web document size) to provide explicit context information to the TCP and IP protocols. While this approach violates the traditional layered Internet protocol architecture, it enables informed decision-making—both at network endpoints and at network routers—regarding flow control, congestion control, and packet discarding. It shows some performance improvement for the web.

3.2. Architectures with Local and Global View of the Network

The above-mentioned cross-layer architectures rely on local information and views, without considering the global networking context or views which may be very useful for wireless networks in optimizing load balancing, routing, energy management, and even some self-behaving properties like self-organisation.

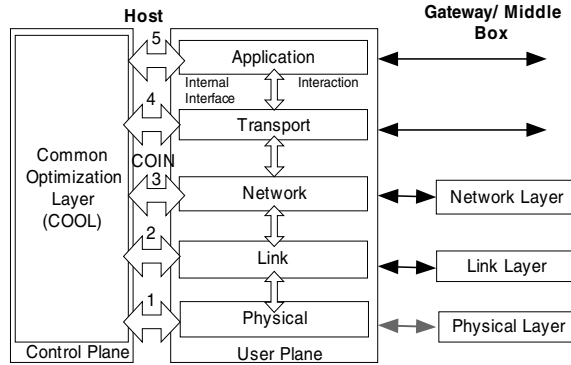


Fig. 9 POEM architecture [8].

Collecting and maintaining network-wide, global statistics can be expensive, while global actions are hard to co-ordinate. However, the effects of such systems can often be dramatic, and they can address problems that are difficult to detect, diagnose or solve using purely local information.

CrossTalk [7] is the only (as far as we are aware) cross-layer architecture (Fig. 10), which has the ability to reliably establish a network-wide, global view of the network under multiple metrics. Having such a global view, a node can use global information for local decision processes in conjunction with a local view containing node-specific information contributed by each layer of the stack or system component. To keep overheads low, no additional messages are sent: instead the local information taken from the local view is piggybacked onto outgoing packets. Piggybacking implies that it is quite unlikely that any node will obtain fully accurate global view under many likely models of data exchange.

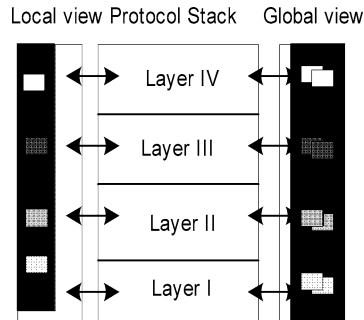


Fig. 10. Cross talk architecture [7].

Even with an uncertain and poor global view, however, CrossTalk has shown performance improvement in a load balancing algorithm specifically reducing per-hop packet delay and bottlenecks. It seems reasonable to expect such performance to be improved by improved global modeling of the network and this expectation is the encouragement for the new cross-layer architecture.

4. CROSS-LAYERING IN AUTONOMIC COMMUNICATIONS

4.1. Self-Behaviors in Autonomic Communications

Autonomic communication is the vision of next-generation networking which will be a self-behaving system with properties such as self-healing, self-configuration, self-organisation, self-optimisation and so forth—the so-called “self-” properties. It can be characterized as service-driven, situated, autonomously controlled, self-organized, distributed, technology independent and scalable. Some of the self-behaviors of autonomic communications are as follows [3]:

Self-awareness: Unpredictable events at run-time often cause systems to drift away from the desired trajectory and behavior, and it is useful that they update their configuration on the fly to enable optimal behaviors in response to any changes. In this way, the autonomic communication system may decide in an *ad hoc* manner which components to remove or to include in the configuration. This requires knowledge about the system’s state (local and global) and resource configuration, as well as a model of the external environment that needs to be extracted, monitored and maintained by the system itself.

Self-organisation: A system is self-organizing if a collection of units coordinate with each other to form a system that adapts to achieve a goal more efficiently without explicit human direction—they form an ensemble rather than a mere collection. Self-organisation can be defined as the emergence of system-wide adaptive structure and functionality from simple local interactions between individual entities. Adaptation involves internalizing information encoded in the functional and non-functional inputs to more efficiently or accurately produce the desired output. The importance of the system’s environment should not be underestimated. Inputs, output, and adaptation are all explicitly dependent on the particulars of the environment the system is in. Self-configuration is perhaps best viewed as an aspect of self-organisation that it restricted to setting interaction parameters for components.

Self-healing: Autonomic communication systems should be conceived with the capabilities to autonomously detect, diagnose and repair localized communication problems resulting from software or hardware failures. In a problem situation, a system component will be attributed the task to report the bug, as well as the

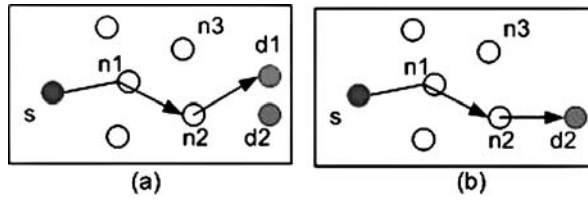


Fig. 11. Self-healing using cross-layering.

source to a manager component. This requires system components to be aware of each other, each other's resources and intended (correct) behaviors.

Self-optimisation: Current and future communication and computing systems have to deal with hundreds of manually set, nonlinear tuning parameters and it makes optimisation task challenging. Autonomic systems will continually seek ways to improve their operation, identifying and seizing opportunities to make themselves more efficient in performance or cost. Just as muscles become stronger through exercise, and the brain modifies its circuitry during learning, autonomic systems will monitor, experiment with, and tune their own parameters and will learn to make appropriate choices about keeping functions or outsourcing them. They will proactively seek to upgrade their function by finding, verifying, and applying the latest updates.

It is evident that to attain these behaviors we require information about the nodes as well as the network. And with the utilization of cross layering approach we can get this information.

4.2. Use of Cross-Layer in Autonomic Communications

It is possible to utilise cross-layer information to attain some of the above self-behaviors. In [8, 21] an effort is made to utilise cross-layer information for self-optimisation and self-healing respectively. Interestingly none of the self behaviors in autonomic computing and communications are extremely orthogonal, which means there is some dependency between them—self-healing is partly supporting self-organisation, and *vice versa*. Following example based on [21] shows the possible use of cross-layering in communication systems in attaining self-healing or self-organizing.

An application scenario of cross-layering in a network of 7 nodes is shown in Fig. 11. In scenario (a) node *s* has a request for node *d1* and it is using the route *s*-*n1*-*n2*-*d1*. Using global view based cross-layer architecture, all the nodes has some knowledge about their direct neighbors, so node *s* has knowledge about *n1*; *n1* has about *n2* and *n3* and so on. If after transmission begins *d1* fails, existing routing protocols would have *n2* receiving the packet, determining *d1* to be dead

and finally sending a “node unreachable” error message to s which wastes all the resources committed to the exchange. Using a cross-layer approach, however, if $d1$ and $d2$ are giving almost same type of services a suitable global view would allow $n2$ to determine that in case of $d1$ ’s failure $d2$ can meet the request of s . This requires making information about the service-level capabilities of a node available to the routing layer, which is facilitated by cross-layering and can easily be expressed as an optimisation algorithm. This leads to scenario (b) where nodes have re-organized because of the death of $d1$, and once $n2$ gets the request from s it reroutes to $d2$ instead of $d1$ and fulfills the request. With this action, cross-layer approach can conserve energy and minimize latency by eliminating the overhead required to invalidate the current route, establish a new route, and retransmit the request. Moreover, it can preserve the original route when failed node becomes available.

5. CONCLUSION AND FUTURE WORK

The worldwide success of the Internet has led to the domination of the layered architecture, but a strict layered design is not flexible enough to cope with the dynamics of next-generation communications which will be dominated by wireless. Careful exploitation of some cross-layer protocol interactions can lead to more efficient performance of the transmission stack (and hence better application layer performances) in different wireless networking scenarios.

In this paper we have described existing cross-layering approaches in next-generation communications. Most of these architectures depend on the local information and only CrossTalk depends on a local as well as a network-wide view to generate a knowledge plane. Even though using an information-gathering process that is quite weak, CrossTalk has shown performance improvements in a load balancing algorithm. We have explored the possible use of cross-layering architecture in autonomic communications. The example of self-healing using cross-layering shows the potential of such approaches to realize the goals of autonomic communications. This also motivates new cross-layer architectures with a hybrid local and global view for autonomic communications, and in our future work we will be developing a cross-layer architecture to better support the collection and application of network-wide views.

ACKNOWLEDGMENTS

This work is partially supported by Science Foundation Ireland under grant number 04/RPI/1544, “Secure and Predictable Pervasive Computing.”

REFERENCES

1. V. Srivastava and M. Motani, Cross-layer design: A survey and the road ahead, *IEEE Communications Magazine*, Vol. 43, No. 12, pp. 112–119, Dec. 2005.
2. J. O. Kephart and D. M. Chess, The vision of the autonomic computing, *Computer*, January , 2003.
3. S. Dobson, S. Denazis, A. Fernndez, D. Gaiti, E. Gelenbe, F. Massacci, P. Nixon, F. Saffre, N. Schmidt, and F. Zambonelli, A survey of autonomic communications, *ACM Transactions on Autonomous and Adaptive Systems*, Vol. 1, No. 2, December 2006. To appear.
4. D. G. Sachs, W. Yuan, C. J. Hughes, A. Harris, S. V. Adve, D. L. Jones, R. H. Kravets, and K. Nahrstedt, GRACE: A hierarchical adaptation framework for saving energy, *ACEED 2005*.
5. D. Kliazovich and F. Granelli, A cross-layer scheme for TCP performance improvement in wireless LANs, *globecom 2004, IEEE Communications Society*, pp. 841–844, 2004.
6. M. Conti, G. Maselli, G. Turi, S. Giordano, Cross layering in mobile Ad hoc network design, *IEEE Computer Society*, pp. 48–51, February 2004.
7. R. Winter, S. Schiller, N. Nikaein, and C. Bonnet, Cross talk: cross-layer decision support based on global knowledge, *IEEE Communications*, Vol. 44, No. 1, January 2006.
8. X. Gu, X. Fu, H. Tshofenig, and L. Wolf, Towards self-optimizing protocol stack for autonomic communication: initial experience, *Proceedings of the 2nd IFIP International Workshop on Autonomic Communication*, Springer Lecture Notes in Computer Science Vol. 3854 (LNCS), pp. 183–201, October, 2005.
9. V. T. Raisinghani and S. Iyer, Cross-layer optimisations in wireless protocol stacks, *Computer Communications*, Vol. 27, No. 2004, pp. 720–724, 2004.
10. S. Dobson, Putting meaning into the network: some semantic issues for the design of autonomic communications systems, *Proceedings of the 1st IFIP Workshop on Autonomic Communications*, volume 3457 of LNCS, Springer Verlag, pp. 207–216, 2005.
11. Z. H. Haas, Design methodologies for adaptive and multimedia networks, *IEEE Communications Magazine*, Vol. 39, No. 11, pp. 106–107, November 2001.
12. G. Wu, Y. Bai, J. Lai and A. Ogielski, Interactions between TCP and RLP in wireless Internet, *Proc. IEEE GLOBECOM'99*, Rio de Janeiro, Brazil, December 1999.
13. K. Chen, S. H. Shan, and K. Nahrstedt, Cross-Layer design for data accessibility in mobile ad hoc networks, *Wireless Personal Communications*, Vol. 21, No. 1, pp. 49–76, April 2002.
14. V. Kawadia and P. R. Kumar, A cautionary perspective on cross-layer design, *IEEE Wireless Communications*, February, pp. 3–11, 2005.
15. G. Wu, Y. Bai, J. Lai, and A. Ogielski, Interactions between TCP and RLP in wireless internet, *Proceedings of IEEE GLOBECOM'99*, Rio de Janeiro, Brazil, Dec. 1999.
16. P. Sudame and B. R. Badrinath, On providing support for protocol adaptation in mobile wireless networks, *Mobile Networks and Applications*, Vol. 6, No. 1, pp. 43–55, Jan. 2001.
17. B. -J. Kim, A network service providing wireless channel information for adaptive mobile applications: Part I: proposal, *Proceedings of ICC'01*, Helsinki, Finland, 2001.
18. Q. Wang and M. A. Abu-Rgheff, A multi-layer mobility management architecture using cross-layer signalling interactions, *Proceeding of 5th European Personal Mobile Communications Conference*, Glasgow, 2003.
19. V. T. Raisinghani and S. Iyer, ECLAIR: An efficient cross layer architecture for wireless protocol stacks, *WWC2004*.
20. Williamson and Q. Wu, A case for context-aware TCP/IP, *Performance Evaluation Review*, Vol. 29, No. 4, pp. 11–23, March 2002.
21. C. Sadler, W. Chen, and L. Kant, Cross-Layer Self-Healing in a Wireless Ad-Hoc Network, U.S. Patent filed April 2005.

Mohammad Abdur Razzaque is a PhD student in the School of Computer Science and Informatics at UCD Dublin. He holds bachelor's and master's degrees from the University of Dhaka in Bangladesh. His research focuses on the autonomic communications.

Simon Dobson is with UCD Dublin's School of Computer Science and Informatics, where his research centres on the semantics of adaptive systems and the programming languages and methodologies used to create them. He has held academic and research posts at Trinity College Dublin and the CLRC Rutherford Appleton Laboratory, and for two years was CEO of a research based start-up company. He has published extensively in the fields of distributed and pervasive systems.

Paddy Nixon is Professor of Distributed Systems in the School of Computer Science and Informatics at UCD Dublin. His interests are in distributed and pervasive computing systems, supported by over 70 internationally reviewed papers. Prior to moving to UCD he was a professor at the University of Strathclyde and academic director of the Kelvin Institute, and academic/industrial partnership funded by the Scottish government.