The Networking Area of the Telecommunication Group at the New University of Lisbon

Paulo Pinto and Luis Bernardo
Faculdade de Ciências e Tecnologia, 2829-516 Caparica, Portugal
Phone: +351-212948545, Fax: +351-212948532, e-mail: {pfp,lflb}@uninova.pt

***Abstract* — This paper describes the main activities of one out of three areas of the Telecommunication Group at the New University of Lisbon – the Networking Area. Our area of intervention is the middleware level of distributed systems. Highly populated systems are the object of research. Mobility and wireless systems is the other main area. The focus in this second area is on management of mobility at middleware and application levels. However, in the effort to build an entire architecture we have proposed a new interworking architecture for future wireless systems. The paper begins with our view of the main driving ideas that will be important in the future systems in our area of intervention. The second part of the paper describes our current research and some near future directions.**

I. INTRODUCTION

The Telecommunication Group of the New University of Lisbon and UNINOVA [1] is divided in three main areas: networking; signal processing; and propagation and radiation. This paper provides an overview of the activities of the networking area of the Group.

II. A THREAD OF ACTION

The current lines of research follow the two main activities the senior members had in the recent past. One of them is Highly Populated Systems, and the other is Internet Technologies. Each line has a follow up but the intersection of both is also a challenging scenario.

Prior to a description of the activities, it might be interesting to understand the motivation that led us to our current work.

A. Highly Populated Systems

If we consider a system with millions of <u>active</u> users and thousands of applications servers, we reach very quickly the conclusion that most of the algorithms we use today will not work [2]. Network protocols and service (and application) structures will also fail. Examples are very easy to find:

- Each time a minor percentage of the users decide to perform a certain task there is congestion. Imagine accessing an on-line journal web site after a catastrophe, of placing a bet in the lotto just a few minutes before the deadline.
- Servers can become so crowded that new approaches have to be taken to solve the problem (for instance, disappear to force users to locate another server).
- The amount of information could be so large that managing a hierarchical structure to hold it can be prohibitive (both in terms of placing new information or updating the existent one).

We have found that with such systems the concept of ***mobility*** just appeared naturally for several reasons [3]: either because new servers have to be placed in the system to raise the response power to users; or because servers cease to exist because they were so overcrowded that it was better for users to find another one.

Another relevant aspect is the management of information. In very large systems, the amount of information can be so large that it becomes impossible to have absolute knowledge of the system for certain issues. The amount of information and the update requirements can be prohibitive. If we cannot have absolute information, then some kind of ***proximity*** must exist. Information starts to be managed by some metrics (distance, age, etc.) and the larger the value of the metric the less users know about it. For instance, when I am driving towards the Lisbon city center it is irrelevant to me how many free places are there in one of Tokyo's parking lots.

This problem of information management is even worse if we let servers to move (or to disappear and appear elsewhere) because updates are hard to perform. Very large information systems nowadays, such as DNS (Domain Name System), work on quasi-static information and partition the information using pathnames.

Highly populated systems can only be feasible (meaning efficient) if we consider a certain level of uncertainty in the algorithms and service structures [4].

This line of research fits nicely in the current problems of ad-hoc and sensor networks.

B. Internet Technologies

One of the successes of the Internet (although many can call it a fault) was to ignore the way users are connected to the Internet. The physical and logical layers were always a dark subject. The important issue was the ability to send and receive IP packets…

A parallel can be drawn with the recent research towards wireless 4G systems. One of the characteristics of 4G systems is the possibility of using one of several radio access networks (RAN) to access the core and the services. Probably the secret in the future 4G architecture might be the irrelevance from the core's point of view of the precise RAN a user is using. Could we adopt the Internet philosophy in 4G system architecture? This idea can shock by being too speculative but there are certain interesting considerations to take into account:

- The Internet has always been a logical, non-real time network until recently. Therefore consistency was far more important than timeliness.
- Addressing and naming in the Internet are confusing issues right from the start. Addresses are network attachment points and a machine can have several. Names are translated using specific services. Mobility is a concept that is finding very hard its way in, also because of this handicap.
- However, it is a characteristic in the Internet/IETF community to solve all the problems at IP level ignoring the lower levels, and Mobile IP was defined to handle mobility.
- Unfortunately, the duration of handovers at Mobile IP level is too long to use this technique for real-time services and applications. We simply cannot make seamless handovers.
- Probably the solution will reside at lower layer but still compatible with all the IP philosophy.

Another interesting thought is the advent of Mobile (Wireless) Internet (MWI). How far from it are we today? Moreover, what characteristics will it have?

The current Internet is assembled with the efforts of Internet providers and users. The core of the network is mainly a (plain) routing engine and users just control most of the services and applications on the network. For instance, the DNS can play a major role in redirections, creation of servers and names, and users can define the DNS database content if they are given permission for that. Another example is the establishment of tunnels, and even the concept of call that is an end-to-end transport layer concept with not intervention of the core network. A last example is the knowledge of the attachment point of the terminal (in terms of Autonomous Systems) via its IP address.

We could think on the MWI as the direct displacement of such a structure. There will be DNS and the only difference is that the access network is radio and not wires. It is very much what we already have now. However, we can also aim at providing users with the same kind of intervention on the network as they do nowadays. The major difference is on the core of the network – there are much more possibilities (due to the Intelligent Networks part of the network) than simply the routing of IP packets. The Internet is also becoming real-time which is another reason to be able to control the access parts. However, even more interestingly is not to stop at this stage. Currently, applications and services see an IP address of an equipment that is apparently wireless and mobile but this is irrelevant because all the mobility and location control is performed by the cellular core without any intervention (or knowledge) of the application. The IP address is stable and fixed at the interface of the GGSN (Gateway GPRS Support Node).

Another consideration is the degree of the penetration of the 2.5G/3G cellular systems and its ubiquity. It is so devastating in the recent years that any mobile and wireless internet solution has to take these systems into consideration. Any facilities that the 4G systems will bring have to be add-ons to the current set of 3G services and user perceptions.

Therefore, to be able to construct the real MWI, the 3G cellular systems must form some basic ground from where the overall system will be defined. Applications and services must take advantage of the terminal location until the utmost piece of information (routing area, cell id, etc.). I.e., the great deal of functionality that belongs to a large untouched cloud inside the core network must be disclosed and made accessible to third party entities to enable them to construct applications that can take advantage of mobility in all its aspects (clients, servers, users, etc.). Access to that information is not dramatic because there are already rules to allow a controlled access – e.g. OSA (Open Services Architecture) [5].

The real MWI will then be a mixture of facilities placed above and outside of the core network (using services such as the DNS, etc.) and facilities placed on OSA – AS (OSA – Application Servers) accessing core's internal information and driving applications with it.

C. Intersection of both areas

The intersection of both areas is a natural follow up. Future MWI will be composed by millions of WPANs (e.g. Bluetooth), WLANS (e.g. 802.11), WMANs (e.g. 802.16), WWAN (e.g. 3G), and fixed nodes, connected though ad hoc networks or directly to a core wired network. It is not difficult to predict that there will be millions of small devices and terminals that can be addressed locally or globally and new concepts must be created in order to construct applications swiftly.

III. How to Pursue

One of the characteristics of the research in middleware subjects is the amount of issues that must be taken into consideration. Very easily a small project can fall into the solution of the universe. Therefore we try to focus very much our research lines in order to propose specific solutions to little problems and not extensive lists of functionalities that make impressive system architectures.

Highly populated ubiquitous computing environment combining ad hoc and infra-structured systems present a challenge not easily addressed by existing approaches, mostly based on transparent usage of the network (Mobile IP hides the network handoffs; grid and object based middleware hide the service location and the access protocol). Optimal solutions depend on the system (e.g. ad hoc networks require different solutions from infra-structured ones) and on the application requirements (e.g. different consistency requirement lead to specific consistency protocols). We argue that such transparency must be optional, and that the application must be able to decide if it uses or not contextual information.

Mobile ad hoc networks (MANETs) are inherently unstable, making it difficult to implement centralised data repositories with name translation information (e.g. DNS, LDAP, etc.) or partial databases with server information (e.g. Distributed Hash Tree based Peer-to-peer algorithms like Chord, CAN, Tapestry, etc.). The information management in this kind of systems is vital since the cost of maintaining references to nodes many hops far away is very high.

Most ad hoc routing protocols adapt to the network conditions by using an on demand approach – they only search a node by its IP address and maintain the path to it, as long as the connection is needed. On demand searching approaches can also be used for arbitrary information, using an unstructured peer-to-peer protocol (e.g. Gnutella). These protocols reduce the information management cost, but they may also produce another kind of problems (poor bandwidth usage, broadcast storm at the MAC layer, etc.). One research line is presently developing a solution to the information problem adapted to MANETs, and to the interconnection of MANETS with infra-structured systems. It extends the highly populated systems scenario, with a mobile active network incorporating the resilience to extreme volatile conditions. The algorithm creates a virtual backbone, composed by the set of stable neighbours, which reduce the overall search cost. A prototype is also being developed by extending JXTA [6] core protocols and services to fit the ad hoc network requirements. This work is waiting for publication. This research line final objective is to propose an integrated solution, which can provide contextual information and a computation environment to the applications using a grid middleware approach.

Quality of service is a key requirement for the MWI. QoS must be supported at the infra-structured network using scalable approaches, and some form of best effort QoS can also be supported by the ad hoc network. For nodes with several active interfaces, QoS can be improved by using the best connection available. Another aspect to be explored is to assess other ways to obtain a certain QoS level. The classical approach is via resource reservation (signalling, reservation protocol, etc.); another approach uses active components within the network (application proxies) that work closely with the applications to perform specific operations (examples are media filtering or conversion, but can also work at semantic level in order to reduce the amount of information). These proxies benefit from better network and computational conditions. Again, the grid middleware can provide a secure and audited support for sharing resources on the ubiquitous computing environment available at the MWI. The middleware may use specific contextual information and resources (e.g. ad hoc status, OSA interfaces, factories, QoS, etc.) to launch components, and make default decisions on the components communication protocols and behaviour. Application components must then be allowed to customise their behaviour, using the information provided by the grid interface.

The second major line of research is related with 4G wireless systems. Handovers inside the 2.5G systems are very fast and seamless to the users. The same is happening in 3G systems, even between different radio technologies. Therefore, it will not be acceptable to have long durations for inter-radio systems handovers in 4G systems. Mobile IP is such a high level approach that even using the several optimizations studied so far one cannot reduce the handover to acceptable limits. We recently proposed an interworking architecture for 4G based on the core network of 3G [7]. The approach, called **core level approach**, is based on the introduction in the core network of a component responsible to manage another (or other) radio access networks. The solution is completely compatible with IP and handover times are below 100 milliseconds. The cellular network forms a primary network for control purposes (paging, location updates, etc.) and also for data, and other radio networks are used on an availability basis. This notion of a primary network is very important because certain radio technologies were not designed to support efficiently certain control procedures (for instance paging). If we try to harmonize these procedures for all radio technologies we could end up slowing down the various processes (e.g., paging operation).

An overview of the necessary signalling messages as well as the authentication and authorization procedures was made and is waiting for publication.

The current work in this area covers the application and services. If the ability of multi-radio connection exists (simultaneously, or not) what do the users want to do with it? One trivial approach is to manage the choice of the connection to use: there is the concept of always best connected (ABC) meaning that one radio access network will be chosen based on some criteria. However, we could think on more interesting scenarios where the application can have a say on the choice of the connection. It can happen that sometimes the concept of ABC could be not connected at all for the data path, and only exchange small control messages for the application in order to wait for cheaper (better, etc.) radio access networks. To be able to construct these kind of applications in an easy way it is important to have powerful (and yet simple) concepts on the way lower levels are used. This is a middleware task: an entity at middleware gather and manage information about the mobile terminal (from the core network and/or from the terminal itself) and offers the concept to the applications.

Most of our current work is based on what is called PS services (Packet switched) of the 3G systems. But there is also an objective to use other radio networks for real-time services. So, the present CS (Circuit switched) services will have to use packets in other radio networks and still comply with the QoS of the circuits in 2G systems. This reality poses very interesting challenges and it will be addressed by us in the near future.

IV. FINAL WRAP UP

A paper such as this does not have properly a conclusion section and the final section is neither the best place to have a critical assessment of the work just described. Therefore, some additional information on the Group will close this paper…

The Telecommunication Group at the New University of Lisbon is a recent group with some work already published widely. Most of the persons came from other institutions and new members are just joining in. The Networking area is working closely with the other two areas and the overall group has strong links to other Sectors in the Electrical Engineering Department (especially with the microelectronics area).

REFERENCES

[1] TG – Telecommunications Group at the New University of Lisbon home page: http://tele1.dee.fct.unl.pt/

[2] L. Bernardo, "A System for Millions of Active Users with Correlated Patterns of Traffic", Instituto Superior Técnico, Lisboa, Portugal, Ph. D. Thesis, 2002 (in Portuguese).

[3] L. Bernardo and P. Pinto, "Scalable Service Deployment using Mobile Agents", in Proc. 2nd International Workshop on Mobile Agents (MA'98), Springer LNCS Vol. 1477, Sep. 1998, pp. 261-272.

[4] L. Bernardo and P. Pinto, "A decentralized Location service Applying P2P technology for picking replicas on replicated services", in Proc. 1st International Conference on E-Business and Telecommunication Networks (ICETE'2004), Setúbal, Portugal, August 2004, Vol.1, pp.39-47.

[5] 3GPP TR 29198-01-620, "3GPP, Technical Specification Group Core Network; Open Service Access (OSA); Application Programming Interface (API); Part 1: Overview (Release 6)", October 2004.

[6] JXTA home page: http://www.jxta.org

[7] P. Pinto, L. Bernardo, and P. Sobral, "UMTS-WLAN Service Integration at Core Network Level", in 3rd European Conference on Universal Multiservice Networks (ECUMN'2004), Porto, Portugal, LNCS, Springer-Verlag Heidelberg, Volume 3262/2004, October 2004, DOI: 10.1007/b101785, pp. 29-39.