Energy efficient wireless ATM design

(Extended abstract)

Paul J.M. Havinga, Gerard J.M. Smit, Martinus Bos
University of Twente, department of Computer Science, the Netherlands
{havinga, smit, m.bos}@cs.utwente.nl

Energy efficiency is an important issue for mobile computers since they must rely on their batteries. We present an architecture for wireless ATM and a novel MAC protocol that achieves a good energy efficiency of the wireless interface of the mobile and provides QoS support for diverse traffic types. The main principles used are to avoid unsuccessful actions, to minimise the number of transitions, and to synchronise the mobile and the base-station. The protocol is able to provide near-optimal energy efficiency (i.e. energy is spent for the actual transfer only) for a mobile within the constraints of the QoS of all connections in a cell, and only requires a small overhead. The architecture of the network interface uses a dynamic error control adapted to the QoS and traffic type of a connection, and has dedicated connection queues and flow control for each connection.

1 Introduction

Portable computers like PDAs and laptops that use wireless communication to interact with the environment rely on their limited battery energy for their operation. Energy consumption is becoming the limiting factor in the amount of functionality that can be placed in these devices. More extensive and continuous use of network services will only aggravate this problem. However, even today, research is still focussed on performance and (low power) circuit design. We believe that it is more effective to save energy by a carefully designed architecture of the mobile, the communications device and wireless communication protocols that consider judicious use of the available energy [9]. Energy reduction should be considered in the whole system of the mobile and through all layers of the protocol stack, including the application layer [1]. Medium access protocols for wireless ATM networks typically address performance metrics such as throughput, efficiency, fairness and packet delay. In this paper we address the additional issue of *energy efficiency*. Considerations of energy efficiency are influenced by the trade-off between energy consumption and achievable Quality of Service (QoS). The characteristics of the wireless channel a MAC protocol has to deal with are basically: high bit error rate, limited bandwidth, broadcast transmission, and high energy consumption. These wireless link particularities and the requirements to provide access for different connection types with a variety of traffic characteristics and QoS requirements, makes the design of a MAC protocol and wireless interface a non-trivial task. It requires a flexible yet simple scheme that is able to adjust itself to different operating conditions in order to satisfy all connections and overall requirements like efficient use of resources such as energy and radio bandwidth.

In this paper we use a novel approach of wireless network design in which various aspects of energy saving approaches are used. The architecture and access protocol is very flexible and provides QoS based on individual connections. It provides a mechanism that allows QoS renegotiations between applications and system. Both the base station and the mobile use the operating system *Inferno* which allows application and system functions to be split and migrated dynamically between client and server.

2 Energy efficient wireless design

The objective of energy efficient wireless ATM design is to meet the required QoS, while minimising the energy consumption of the *mobile*. These requirements often conflict, and a trade-off has to be made. We define *energy efficiency* as the quotient between the intrinsic amount of energy needed to transfer a certain quantity of data and the actually used amount of energy (thus including all overheads).

2.1 Sources of unessential energy consumption

The focus of this work is on minimising the energy consumption of a mobile and in particular the wireless interface, the transceiver. Several causes for unessential energy consumption exist:

- Many applications have little demanding traffic needs, and the transceiver is *idling* most of the time. Measurements show that on typical applications like a web-browser or e-mail, the energy consumed while the interface is on and idle is more than the cost of actually receiving packets.
- The typical *inactivity threshold*, which is the time before a transceiver will go in the off or standby state after a period of inactivity, causes the receiver to be in an energy consuming mode needlessly for a significant time.
- In a typical wireless broadcast environment, the receiver has to be powered on at all times to be able to *receive messages* from the base station, resulting in significant energy consumption.
- Significant time and energy is further spent by the mobile in switching from transmit to receive modes, and vice-versa. The *transition* between these modes typically takes between 6 to 30 μs. The transition

from sleep to transmit or receive generally takes even more time (250 µs).

- Another main cause of energy consumption are the *collisions* that occur mainly at high load situations. Collision causes all parties involved to retransmit the original data, leading to unnecessary energy consumption and also to possible unbounded delays.
- The *overhead of a protocol* also influences the energy requirements due to the amount of 'useless' control data and the required computation for protocol handling. In general, simple protocols need relatively less energy than complex protocols.
- The high *error rate* that is typical for wireless links might cause another source of energy consumption for several reasons. First, when the data is not correctly received the energy that was needed to transport and process that data is spoiled. Secondly, energy is used for error control mechanisms. Finally, because in wireless communication the error rate varies dynamically over time and space, a fixed-point error control mechanism that is designed to be able to correct errors that hardly occur, spoils energy and bandwidth. If the application is error-resilient, trying to withstand all possible errors spoils even more energy for needless error control.

2.2 Principles of energy efficient MAC design

From the above observations we can deduce the following three main principles to design a MAC protocol that is energy efficient for the mobile.

- *Avoid unsuccessful actions of the transceiver* -- Two main topics cause unsuccessful actions: collisions and errors. Every time a *collision* occurs energy is spoiled because the same transfer has to be tried again. A protocol, in which a base-station broadcasts a traffic control for all mobiles in range with information about when a mobile is allowed to transmit or is supposed to receive data, reduces the occurrence of collisions significantly, and will be more energy efficient.
  Errors on the wireless link can be overcome by mechanisms like retransmissions or error correcting codes. The error control mechanisms can be adapted to the current error condition in such a way that it minimises the energy consumption needed and still provides (just) enough fault tolerance for a certain connection. A different strategy to reduce the effect of errors is to avoid traffic during periods of bad error conditions.
- *Minimise the number of transitions* -- Scheduling traffic into bursts in which a mobile can continuously transmit or receive data can reduce the number of transitions. The number of transitions needed can also be reduced by collecting multiple requests of multiple applications on a mobile, and by piggy-backing new requests on current data streams. Simple protocols can further reduce the required amount of transitions due to the low amount of control messages needed.
- *Synchronise the mobile and the base-station* -- When the base-station and mobile are synchronised in time, the mobile can go in standby or off mode, and wake up just in time to communicate with the base-station. The energy consumption needed for downlink traffic can be reduced when the time that the receiver has to be on - just to listen whether the base-station has some data for the mobile - can be minimised. If the wake-up call of the communication is implemented with a low-power low-performance radio, instead of the high-performance high-energy consuming radio, then the required energy can be reduced even more.

Note that these principles can reduce the energy consumption of the wireless interface. The energy consumption of the mobile system is much more complex and comprises many issues. The total achieved energy reduction is thus based on many trade-offs. For example, grouping traffic in multimedia video streams to minimise the number of transitions, requires data to be buffered in the client's memory. The required amount of energy needed for buffering reduces the effect of the energy savings principle in some sense.

3 E^2MaC: energy efficient medium access control

This section describes the basic principles and mechanisms of our energy efficient medium access control for wireless links, called E^2MaC. The protocol and architecture is targeted to a system in which quality of service and energy consumption plays a crucial role. A typical example of such a system is the *Mobile Digital Companion* that is investigated in the Moby Dick project [8]. The structure is composed of several base-stations that each handle a single radio cell covering several mobile stations. Each mobile can have multiple unidirectional connections with different Quality of Service. They can have the following service classes defined by the ATM Forum traffic management group: constant bit rate (CBR), real time VBR, non-real time VBR, unspecified bit rate (UBR), and available bit rate (ABR).

3.1 The Mobile Digital Companion

The Moby Dick project develops and defines the architecture of a new generation of hand-held computers, the so-called Mobile Digital Companion. It is a small personal portable computer and wireless communications device that can replace cash, cheque book, passport, keys, diary, phone, pager, maps and possibly briefcases as well. The energy consumption due to the increasing demand for performance and functionality will be the limiting factor for its capabilities. Therefore reducing energy consumption plays a crucial role in the architecture.

The Mobile Digital Companion has a rather unconventional architecture that saves energy by using system decomposition at different levels of the architecture and exploiting locality of reference with dedicated, optimised modules [4]. The approach is based on dedicated functionality and the extensive use of power reduction techniques at all levels of system design. The system has an architecture with a general-purpose processor surrounded by a set of heterogeneous *autonomous* programmable modules, each providing an energy efficient implementation of dedicated tasks. A reconfigurable internal communication network switch exploits locality of reference and eliminates wasteful data copies. The switch is a simplified ATM switch as it is capable to transfer data streams consisting of ATM packets to their destination according their virtual connection identifier.

One of the functional modules of a Mobile Digital Companion is the network module. This module provides the interface between the external world and the different modules of the system. The main processor is responsible for the establishment of the connections between the modules, but also negotiates with the external infrastructure about the QoS of the connections between network module and the modules that are at the end-point of connections. Once a connection between modules is established, they autonomously communicate via the switch.

3.2 E^2MaC architecture

The MAC and data link layers are the main functions to be performed by the network module. This basic functionality is assigned to several functional entities. Figure 1 depicts the E^2MaC architecture. It does not show the QoS management that is located in the base station only. The number of connection queues in the figure is just an example.

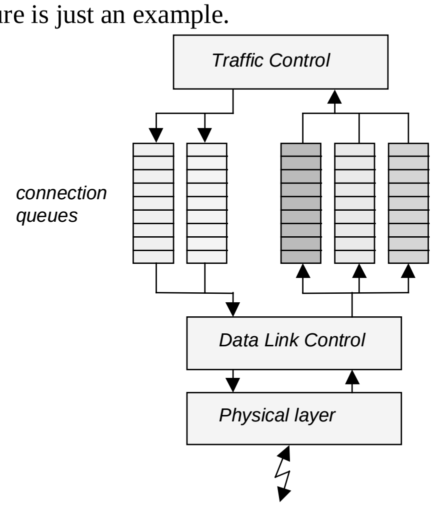

Figure 1: E^2MaC architecture.

Data Link Control (DLC) performs the traffic allocation of data in the transmission queues and forwards data from the reception queues. The actual admission decision of connections is made by the QoS management, which informs the Data Link Control using a traffic control packet (either transmitted over the air for the mobile or internally for the base-station). Data Link Control regulates the flow of ATM cells between the physical layer and local *buffer*. This buffer is only meant to store ATM cells for a short time, just enough to implement an effective error control mechanism. The buffer is organised in such a way that it has a small queue for each connection. The Data Link Control performs error detection and basic error correction on each ATM cell. The overhead required for error control is fixed, so that the slot size does not vary.

The *Traffic Control (TC)* controls the flow of data from the connection queues to the corresponding end-points and applies an adaptive error control scheme that operates on individual virtual connections. Flow control is needed to prevent buffer overflow. ATM cells of a connection on which the maximum allowed delay is exceeded, for example due to bad error conditions, will be discarded by traffic control.

QoS management is located in the base-station only. It establishes, maintains and releases wireless connections between the base-station and the mobile. It is the intermediate between the wired and the wireless world, and negotiates about the QoS traffic contract of a connection in lieu of the mobile. This entity also provides support for handover and mobility services. The QoS management also contains the *slot scheduler* that assigns bandwidth to active connections. This schedule (called traffic control) is broadcasted to all mobiles so that they know when they are allowed to transmit, or are expected to receive data. In composing the traffic control, the slot scheduler takes into account: the state of the downlink and uplink queues, the radio link conditions per connection, and spare capacity for handovers.

3.3 Frame structure

The protocol uses fixed-length frames of multiple slots. The base-station and mobile are completely synchronised, which allows the mobile to power-on precisely when needed. A frame is divided in slots that can have three basic types: *traffic control*, *registration request*, and *data*. The base-station controls the traffic for all mobiles in range of the cell and broadcasts the schedule in the traffic control slot. The data slots are used to send the actual data. The traffic of a mobile is grouped as much as possible into packets. New reservations or updates of current reservations for the next frames can be piggy-backed onto a packet. Only new connections may encounter collisions, the data slots are collision-less. All remaining slots can be used by the mobile to make *connection requests.*

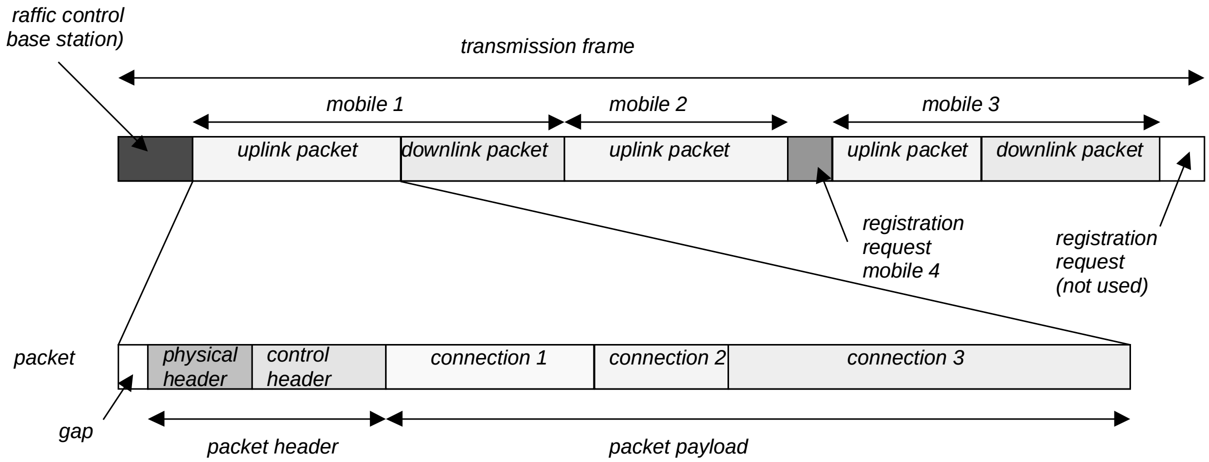

Figure 2: Example of a transmission frame.

3.4 MAC and data link functions

The general theme that influences many aspects of the design of the MAC protocol is adaptability and flexibility. This implies that for each connection a different set of parameters concerning scheduling, flow control and error control should be applied.

Dynamic slot scheduling -- The QoS manager of the base-station schedules all traffic according to the QoS requirements and tries to minimise the energy consumption of the mobiles by minimising the number of transitions the mobile has to make. It therefor schedules the traffic of a mobile such that all downlink and uplink connections are grouped as much as possible into packets within the limitations imposed by the QoS of the connections. The uplink and downlink packets for a mobile are also grouped sequentially (if possible) so that the mobile can power down for longer period and makes minimal transitions between power modes. The slot scheduler has to dynamically adapt its schedule when 1) connections are added or removed, 2) connections change their QoS requirements, and 3) the channel between mobile and base station has significant change in error condition. The scheduler can try to avoid periods of bad error conditions by not scheduling non-time critical traffic during these periods.

Dynamic flow control -- The connection queues of the connections can have a different size and replacement policy. The Data Link Control of both the mobile and the base station however just need to know whether it can sent data for a connection or not. The simplest implementation of flow control therefor consists of a stop/run mechanism. The flow control information is transmitted in the header of each data packet. This mechanism allows the slot scheduler to adapt the scheduling and also allows a mobile to inform the base-station that it is ready to receive downlink traffic for a certain connection. This message is also needed when the mobile wakes up and notices that the base station has traffic waiting in its queue.

Dynamic error control -- Due to the dynamic nature of wireless networks, adaptive error control can give significant gains in bandwidth and energy efficiency [6]. Adaptation of the error control can be influenced by three considerations: 1) the FEC redundancy can be adapted to the channel bit error rate and induced energy consumption, 2) the error control algorithm can be adapted to the required quality, 3) the performance of various error-correcting methods depend on the actual error statistics of the transmission channel. The selection of the error control scheme and the required size of the queue depend on the QoS constraints imposed on each connection, such as delay constraints or loss-less transfer constraints. This avoids applying error control overhead to connections that do not need it, and allows the possibility to apply it selectively to match the required QoS and the conditions of the radio link. The error control will be based on adaptive error *correcting* techniques. Although good designed retransmission schemes can be energy efficient, they are much more complex to implement (they require a protocol with control messages, sequence numbers, retry counters, etc.) and can introduce intolerable low performance in delay, jitter and bandwidth to fulfil the required QoS of the connection [3]. The redundant data needed to implement the error correction are multiples of ATM cells, so that they fit well in a transmission frame. The Traffic Control will try to match the radio conditions to the applied fault tolerance, and adapt the required bandwidth accordingly.

Above these error control adaptations, the slot scheduler can also adapt its scheduling policy to the error conditions of wireless connections to a mobile. The scheduler can try to *avoid periods of bad error conditions* by not scheduling non-time critical traffic during these periods. Note that the error conditions perceived by each mobile in a cell could be different. Since the base station keeps track of the error conditions per connection (and thus also per mobile), it can give other mobiles more bandwidth when these have better conditions. This can lead to a throughput that may even exceed the average rate on the channel, due to the introduced dependence between admitted connections and channel quality [2].

3.5 Future design issues

As discussed in section 2.1, a main source of unessential energy consumption is due to the costs of just being connected to the network. In E^2MaC we have tried to minimise this overhead, but it still requires the receiver to be switched on from time to time, just to listen whether the base-station has some messages waiting. Therefor we consider using a low-power low-bandwidth receiver for the signalling only. This receiver will be used to wake-up a mobile by the base-station. It uses the same synchronisation mechanism between mobile and base-station, but now uses a simple, low performance, low power receiver.

The energy per bit transmitted or received tends to be lower at higher bit rates. For example, the WaveLAN radio operates at 2Mb/s and consumes 1.8 W, or 0.9 µJ/bit. A commercially available FM transceiver (Radiometrix BIM-433) operates at 40 kb/s and consumes 60 mW, or 1.5 µJ/bit. This makes a low bit rate radio less efficient in energy consumption for the same amount of data. However, when a mobile has to listen for a longer period for a broadcast or wake-up from the base station, then the high bit rate radio will consume about 30 times more energy than the low bit rate radio. Therefor, the low bit rate radio must only be used for the basic signalling, and as little as possible for data transfer.

4 Conclusions

We have presented a novel architecture for an ATM network interface and a MAC protocol that provides support for diverse traffic types and QoS while achieving a good energy efficiency of the wireless interface of the mobile.

The main complexity is moved from the mobile to the base station that has plenty of energy. The scheduler of the base station is responsible for providing the required QoS for the connections on the wireless link and tries to minimise the amount of energy spend by the mobile. The main principles of the E^2MaC protocol are: avoid unsuccessful actions, minimise the number of transitions, and synchronise the mobile and the base-station. The protocol is able to provide near-optimal energy efficiency (i.e. energy is spent for the actual transfer only) for a mobile within the constraints of the QoS of all connections.

In the architecture we have minimised the energy consumption by using a flexible slot scheduler that minimises the energy consumption of the mobile, while still maintaining the QoS requirements for the connections. The error control and buffer management further optimise the system for both performance and energy consumption. Each connection has its own adaptable error control and dedicated connection queues with dynamic flow control. QoS renegotiations may be required to assure a lower, but feasible level of service. Many multimedia applications can deal with varying bandwidth availability once provided with sufficient knowledge about the operating conditions.

Currently we are implementing the protocol and building a prototype. Both base stations and mobiles use the operating system *Inferno* from Lucent Technologies Bell Laboratories. The operating system allows application- and system functions to be split easily. Certain functions of the system can be migrated from the portable system to a remote server. For example, parts of the communications protocol are processed on the base station in lieu of the mobile. The mobiles can thus use a dedicated lightweight protocol to communicate with the base station rather than for example TCP/IP or UDP. The MAC layer protocol will be implemented in a low-end micro-controller with some external memory and programmable logic (FPGA). The physical layer is based on the WaveMODEM module that is used to build the WaveLAN system.

5 References

[1] Chen, et al. “Comparison of MAC Protocols for Wireless Local Networks Based on Battery Power Consumption”, *IEEE Infocom’98*, San Francisco, USA, pp. 150-157, March 1998.

[2] Chockalingam, A., Zorzi, M.: “Energy consumption performance of a class of access protocols for mobile data networks”, *VTC’98*, Ottawa, Canada, May 1998.

[3] Havinga, P.J.M., “Energy efficiency of error correcting mechanisms for wireless communication”, *CTIT technical report* 98-19, 1998, the Netherlands.

[4] Havinga P.J.M., Smit G.J.M.: “The Pocket Companion's Architecture”, *Euromicro Summer School on Mobile Computing '98*, pp. 25-34, Oulu, Finland, August 1998.

[5] Mangione-Smith, B. et al.: “A low power architecture for wireless multimedia systems: lessons learned from building a power hog”, *ISLPED 1996*, Monterey CA, USA, pp. 23-28, 1996.

[6] Schuler C.: "Optimization and adaptation of error control algorithms for wireless ATM”, *International Journal of Wireless Information Networks*, Vol. 5, No. 2, April 1998.

[7] Sivalingam, K.M., Srivastava, M.B. Agrawal, P.: “Low power link and access protocols for wireless multimedia networks”, *Proceedings IEEE Vehicular Technology Conference*, Phoenix, AZ, pp. 1331-1335, May 1997.

[8] Smit G.J.M., et al.: “Overview of the Moby Dick project”, *Euromicro Summer School on Mobile Computing '98*, pp. 159-168, Oulu, Finland, August 1998.

[9] Woesner H., Ebert J., Schläger M., Wolisz A.: "Power-saving mechanisms in emerging standards for wireless LANs: The MAC level perspective", *IEEE Personal Communications*, Vol. 5, No. 3, June 1998.