Service Differentiation Extensions for IEEE 802.11

Albert Banchs[ab], Xavier Perez[ab], Markus Radimirsch[c], Sebastia Sallent[b1]

[a] C&C Research Laboratories, NEC Europe Ltd., Heidelberg, Germany

[c] Departament de Telematica, Universitat Politecnica de Catalunya, Spain

[c] Institut für Allgemeine Nachrichtentechnik, University of Hannover, Germany

The world of data communication has undergone many changes over the last few years. Probably the most important one is the convergence of voice, video and data communication under the roof of the Internet Protocol (IP) suite. Originally, IP was designed to support elastic services, i.e. data applications like file transfer, electronic mail and remote terminal. Elastic services are tolerant of delays and, even though they benefit from increasing data rates in terms of user satisfaction, still work at low data rates. Voice services, in contrast, require a certain minimum rate and suffer significantly from high delay and delay variation. In wired networks, the delay issue can be solved by (1) bandwidth overprovisioning and (2) service differentiation. Bandwidth overprovisioning is not possible in radio networks and, thus, innovative solutions for service differentiation are necessary. This contribution addresses this issue by extending the MAC protocol of the IEEE 802.11 wireless LAN standard.

The extension is divided in two steps. In the first step, real-time services are distinguished from elastic services by a priority scheduling approach in order to meet the hard requirements of e.g. voice communication. In the second step, service differentiation is introduced for elastic services. Elastic service differentiation can be on a per-service basis, e.g. to differentiate between an interactive service like web access and a non-interactive one like electronic mail, on a per-user basis, e.g. to differentiate services for managers and other employees in a company, or even on a per-device basis, e.g. for different sensors in a production facility. The fact that the two-step extension proposed is designed in different and independent modules gives the manufacturer the option to omit one of them, therewith simplifying the migration effort from the current IEEE 802.11 standard.

The real-time extension is a distributed scheme which redefines the Point Coordination Function (PCF) of the current IEEE 802.11 standard. The original PCF is not widely supported in current products, and the only requirement of our solution is that the original PCF must not be used in a network together with the extension presented here. We propose a scheme to resolve contention among real-time stations based on two Elimination Bursts. This scheme provides a constant residual collision rate almost independent from the number of contending stations. Admission control is a key aspect of the real-time extension. In order to meet the requirement of real-time traffic for low delay, the amount of traffic using this service must be kept sufficiently low. In addition, admission control for real-time traffic avoids the starvation of lower priority traffic.

Applying admission control, however, does not match the nature of elastic services. Elastic services do not need a specific capacity but rather as much capacity as possible. Therefore, we propose a relative differentiation model in which a high priority service always receives a higher throughput than a low priority one. Due to the relative nature of the scheme, admission control can be omitted and differentiation can always be achieved, independent of the incoming load.

The elastic service differentiation model is based on the notion of a *share*. The *share* reflects the priority assigned to a service such that the throughput experienced by the service is proportional to the *share* assigned to it. As an example, assume eight services, six with a *share* of 1 and two with a *share* of 2, then the total of shares is 10. Each of the two high priority services should get 20% of the total throughput and the remaining six should get 10% each.

We propose an algorithm that dynamically computes the size of the Contention Window (CW) of the IEEE 802.11 DCF mode. With this algorithm, each station computes its relative throughput W which is equal to its measured transmission rate divided by the *share* assigned to its service. This value is included in the header of each transmitted packet. Whenever a station observes a packet from another station with a W different from its own W, this station modifies its own CW by a small amount such that the difference between its own W and the observed W is reduced. In this way, the W of all the stations converge toward a common value and the desired differentiation is achieved.

[1] This work has been partially funded by CICYT with the project TIC-98-0495-C02-01.

A mechanism to cope with overload situations is introduced. Note that the proposed algorithm is backwards compatible with existing 802.11 products.

The performance of the proposed two-step architecture has been extensively evaluated through simulation. The inverse delay distribution function and packet drops for real-time traffic have been studied for a varying number of real-time stations with different source rates. Simulation results show, for example, that up to 6 real-time stations with a transmission rate of 128 Kbps can be supported with delay of less than 25 ms for more than 97% of the packets. Elastic traffic has been simulated as a function of the shares and the total number of stations for constant bit rate, bursty and TCP sources. Simulation results show for all cases a good behaviour in the achieved throughput differentiation, the efficiency of the overall utilisation of the channel and the percentage of packet drops.

Service Differentiation Extensions for IEEE 802.11

Albert Banchs[ab], Xavier Perez[ab], Markus Radimirsch[c], Sebastia Sallent[b]

[a]C&C Research Laboratories, NEC Europe Ltd., Heidelberg, Germany

[b]Departament de Telematica, Universitat Politecnica de Catalunya, Spain

[c]Institute für Allgemeine Nachrichttechnik, University of Hannover, Germany

© NEC Europe Ltd., NDL-E Heidelberg

Introduction

- One of the most important changes in the world of data communications in the last few years is the convergence of voice video and data under the roof of the Internet Protocol suite.
- Originally, IP was designed to support elastic services, i.e. data applications like file transfer, electronic mail or remote terminal.
- These elastic services are tolerant of delays.
- In contrast, real-time services like voice suffer significantly from high delay and delay variation.

© NEC Europe Ltd., NDL-E Heidelberg

Motivation

- In wired networks, the delay issue can be solved by:
  1) Bandwidth overprovisioning
  2) Service Differentiation
- Bandwidth overprovisioning is not possible in radio networks and, thus, innovative solutions for service differentiation are necessary.
- Our contribution addresses this issue by extending the MAC protocol of the IEEE 802.11 Wireless LAN standard.
- We propose two independent extensions:
  - One extension to support real-time traffic.
  - One extensions to provide elastic traffic with service differentiation.

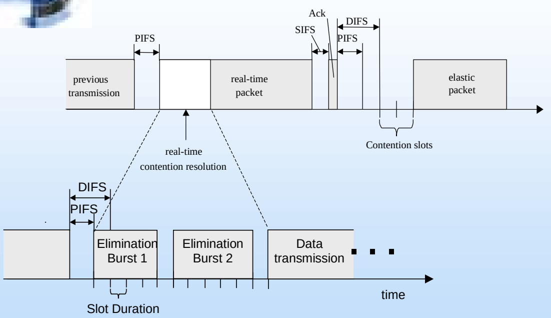
© NEC Europe Ltd., NDL-E Heidelberg

Real-time traffic extension

- Real-time packets require low delays
- Low delays can be provided by:
  - Giving real-time a higher priority than elastic services
  - Limiting the amount of real-time traffic via admission control
- PCF receives a higher priority treatment than DCF by using a smaller IFS.
- In our proposal we redefine the PCF function by allowing Real-time traffic to access the channel after the PIFS while stations with elastic traffic have to wait until the end of the DIFS.
- A contention resolution algorithm is still needed to avoid collisions between stations with Real-time traffic.

© NEC Europe Ltd., NDL-E Heidelberg

Real-time traffic Extension



© NEC Europe Ltd., NDL-E Heidelberg

Elastic Service Extension

- This extension introduces differentiation for elastic services.
- Elastic service differentiation can be on a:
  - per-service basis, e.g. to differentiate between an interactive service like web and a non-interactive one like electronic mail.
  - per-user basis, e.g. to differentiate services for managers and other employees in a company.
  - per-device basis, e.g. for different sensors in a production facility.
- Elastic services do not need a specific capacity but rather as much capacity as possible.
- We propose a relative differentiation model that matches the nature of elastic traffic.

© NEC Europe Ltd., NDL-E Heidelberg

Elastic Service Extension: Relative Differentiation

- Our relative differentiation model is based on the notion of a share.
- The share reflects the priority assigned to a service such that the throughput experienced by the service is proportional to the share assigned to it:

$$\frac{r_i}{S_i} = \frac{r_j}{S_j}$$

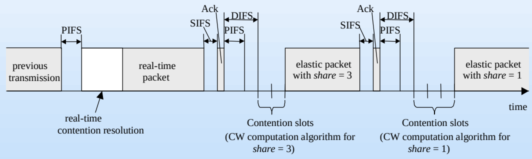
r_i - sending rate of user *i*
S_i - share assigned to user *i*

- As an example, if we have 8 services, 6 with a share of 1 and 2 with a share of 2, each of the two high-priority services should get 20% of the total throughput, and the remaining 6 should get 10% each.
- Note that, in contrast to the real-time extension, this extension does not require admission control.

© NEC Europe Ltd., NDL-E Heidelberg

Elastic Service extension: Contention Window Computation

- We propose an algorithm that dynamically computes the size of the CW of the IEEE 802.11 DCF mode in order to achieve the desired relative differentiation.



© NEC Europe Ltd., NDL-E Heidelberg

Elastic Service extension: Contention Window Computation

- Each station computes its relative throughput *W*:

$$W_i = \frac{r_i}{S_i}$$

- *W* is included in the header of each transmitted packet.
- Whenever a station observes a packet from another station different than its own *W*, this station modifies its CW by a small amount such that the difference between its own *W* and the observed *W* is reduced.
- In this way, the *W* of all stations converge to a common value and the desired differentiation is achieved.
- A mechanism to cope with overload situations is introduced.
- The proposed algorithm is backwards compatible.

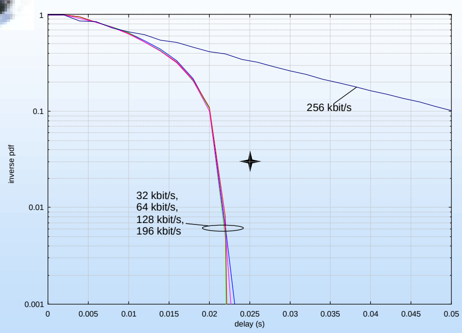
© NEC Europe Ltd., NDL-E Heidelberg

Simulations for the Real-time extension

- Setup:
  - Quality criterion: Max. Delay of 25 ms not exceeded by 97% of the packets
  - Packet length: 500 bytes
  - Total Number of stations: 20
  - 2 Mbps W-LAN
  - All stations have CBR traffic. Stations with elastic traffic always have a packet to transmit
- Results:
  - Saturation throughput approx. 1300 kbit/s
  - Stations with real-time traffic get what they request as long as their total requested data rate stays below the saturation throughput
  - The stations with elastic traffic share the remaining data rate

© NEC Europe Ltd., NDL-E Heidelberg

6 Real-time stations



© NEC Europe Ltd., NDL-E Heidelberg

Observations

- The contention resolution scheme with two elimination bursts can meet the quality criterion defined for up to six stations with data rates up to 196 kbit/s for a 2 Mbps Wireless LAN
- The delay distribution is very steep and has excellent properties
- If the Real-time Stations use in total more than the saturation throughput, the delays increase dramatically
  - In this case, the service quality for all other stations drops drastically

© NEC Europe Ltd., NDL-E Heidelberg

Admission Control for Real-time traffic

No. of stations	Data rate (kbit/s) 32	64	128	256	512
2	x	x	x	x	x
4	x	x	x	x	
6	x	x	x		
8	x				
10					

- As a rule of thumb, the admission control for Real-time traffic could allow not more than 6 stations with a maximum data rate of 128 kbit/s.

© NEC Europe Ltd., NDL-E Heidelberg

Simulations for the Elastic differentiation extension

We studied via simulation the following aspects of our elastic differentiation extension:

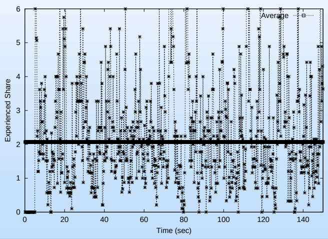
- Differentation experienced
- Channel utilization
- Packet drops
- Backwards compatibility
- Algorithm to avoid overload
- Traffic types: CBR, ON/OFF, TCP

The algorithm proved to behave well for all scenarios.

© NEC Europe Ltd., NDL-E Heidelberg

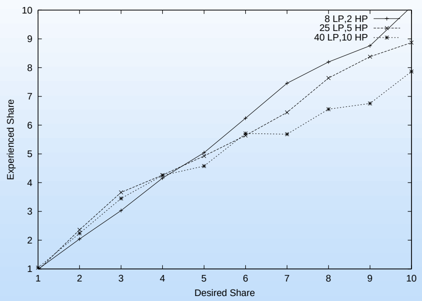
CW algorithm

- The relative differentiation is achieved by adjusting adaptively the CW. In the Figure it can be seen that in average a station with a share equal to 2 gets twice the throughput of a station with a share equal to 1.



© NEC Europe Ltd., NDL-E Heidelberg

Differentiation experienced



© NEC Europe Ltd., NDL-E Heidelberg

Conclusions

- The proposed architecture provides QoS support in Wireless LAN.
- It is based on the IEEE 802.11 standard and it has been designed with the goal of minimizing the migration effort.
- The architecture presented consists of two extensions: one for real-time traffic and the other for elastic traffic.
- The two extensions are adapted to the different natures of the two traffic types, providing absolute performance levels for real-time traffic and relative performance levels for elastic traffic.
- The real-time extension redefines the PCF scheme into a distributed scheme. We argue that distributed control is more efficient and flexible than centralized control.
- The elastic extension modifies the CW computation of the DCF mode of the standard. This modification has been done such that backwards compatibility is provided.

© NEC Europe Ltd., NDL-E Heidelberg