Service Differentiation and QoS in WLANs (802.11e)

Renato Lo Cigno
www.disi.unitn.it/locigno/didattica/NC/
Quest’opera è protetta dalla licenza Creative Commons NoDerivs-NonCommercial. Per vedere una copia di questa licenza, consultare:
http://creativecommons.org/licenses/nd-nc/1.0/
oppure inviare una lettera a:
Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.

This work is licensed under the Creative Commons NoDerivs-NonCommercial License. To view a copy of this license, visit:
http://creativecommons.org/licenses/nd-nc/1.0/
or send a letter to
Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.
What’s the Problem in PCF

• PCF designed to provide QoS to real-time traffic
• What makes QoS in 802.11 difficult?

1. Unpredictable beacon delay
   - A WSTA stops all timers at TBTT thus it does not initiate a transmission after TBTT; however, it continues on-going transmissions, hence beacon may be delayed
   - The larger the frame size, the longer the delay (up to 4.9 ms)

2. Unknown transmission duration

3. Static Polling List
Quality-of-Service Provisioning: Some Terminology

• Definition: A flow is a packet stream from a source to a destination, belonging to the same application

• Definition: QoS is a set of service requirements to be met by the network while transporting a flow

• Typical QoS metrics include: available bandwidth, packet loss rate, estimated delay, packet jitter, hop count and path reliability
QoS in Wireless Networks

• QoS schemes in wired networks are NOT suitable for wireless networks
  - e.g., current wired-QoS routing algorithms require accurate link state and topology information
  - time-varying capacity of wireless links, limited resources and node mobility make maintaining accurate information difficult

• Supporting QoS in wireless networks is an even more difficult challenge
Service Differentiation MAC Schemes that lead to 802.11e

Service differentiation-based MAC schemes

- Station-based
  - DCF-based
    - AC scheme
    - DFS
    - VMAC
    - Blackburst
    - DC scheme
  - PCF-based
    - Priority-based
    - Distributed TDM

- Queue-based
  - DCF-based
    - Per-flow scheme
    - 802.11e EDCF
    - AEDCF
  - PCF-based
    - 802.11e HCF

Thanks: TLC Networks Group – Politecnico di Torino

Nomadic Communications: 802.11e

Renato.Locigno@dit.unitn.it
A QoS Standard for WLANs: IEEE 802.11e

• The IEEE 802.11 TG E was formed in 1999
• The Project Authorization Request (PAR) was approved in March 2000
• **Scopes of the IEEE 802.11 Task Group E**
  - Enhance the current 802.11 MAC to improve and manage QoS
  - Consider efficiency enhancements in the areas of DCF and PCF
  - Provide different classes of service (8 TCs)
802.11e Standard

• Released last autumn
• PHY unchanged (use a/b/g)
• MAC Enhanced: Goals
  - Traffic Differentiation and Guarantee
  - TSPEC and CAC
  - Interoperation with legacy 802.11
802.11e: QSTA, QAP, QBSS, HCF

• A station using 802.11e is called **QoS Enhanced Station** (QSTA)
• An AP using 802.11e is called **QoS Access Point** (QAP)
• QSTA e QAP works within a **QoS Basic Service Set** (QBSS)
• The two coordination functions DCF e PCF are substituted by a single **Hybrid Coordination Function** (HCF)
TXOPs

- **TXOP**: Transmission Opportunity
  - Time interval during which a QSTA has the right to transmit
  - It is characterized by a starting time and a maximum duration (TXOP_Limit)
  - Used in both CP and CFP
Hybrid Coordination Function, alternates:

- EDCA (Enhanced Distributed Channel Access), contention based, conceived to support legacy stations and provide some *stochastic* level of differentiation.

- HCCA (HCF Coordinated Channel Access), polling based, provides collision free periods with guaranteed assignment and *deterministic* differentiation.
802.11e QoS Mechanisms

802.11e proposes a new access scheme: **Hybrid Coordination Function (HCF)**, composed of two coordination functions

- **Enhanced Distributed Channel Access (EDCA)**
  - A basis layer of 802.11e; operates in CP

- **HCF Controlled Channel Access (HCCA)**
  - HCCA operates in CFP
802.11e QoS Mechanisms

- **MAC-level FEC (Hybrid I and II)**
- **Ad hoc features:**
  - Direct Communication / Side Traffic
  - WARP: Wireless Address Resolution Protocol
  - AP mobility
802.11e: Hybrid Coordinator

• Within a QBSS a centralized controller is needed to coordinated all QSTAs. This is the Hybrid Coordinator (HC), normally implemented within a QAP

• An HC has the role of splitting the transmission superframe in two phases continuously alternatnating:
  - *Contention Period* (CP), where QSTAs content for the channel using EDCA
  - *Contention-Free Period* (CFP), where HC defines who is going to use the channel and for what time with a collision free polling protocol
MAC 802.11e: HCCA

Beacon Interval (BI)

- Beacon
- CFP
- CAP
- CP
- CAP
- Beacon

- EDCA
- HCCA
802.11e: EDCF

- The *Enhanced Distributed Coordination Function* (EDCF) define a differentiated access scheme based on an improved (yet complex) contention scheme.
- It is an evolution of CSMA/CA DCF, with the add-on of traffic classes to support QoS and differentiate traffic.
- EDCF is designed to support frames with the same 8 priority levels of 802.1d, but mapping them on only 4 access categories.
- Every frame passed to the MAC layer from above, must have a priority identifier (from 0 to 7), called *Traffic Category Identification* (TCId).
802.11e: EDCF

• TCId is written in one header field of the MAC frame
• Each 802.11e QSTA & QAP MUST have four separated AC queues
• Each AC queue is FIFO and behaves independently from the others as far as the CSMA/CA MAC protocol is concerned
802.11e: EDCF

The diagram illustrates the prioritization and scheduling of traffic classes (AC) for different classes of service (COS). The prioritization is indicated by the arrows pointing from AC 7 to AC 1, with AC 7 having the highest priority. Each AC level contains parameters such as Backoff AIFS, CWmin, and CWmax, which are used to control the access to the channel and prevent collisions. The Virtual Collision Handler is responsible for managing the channel access when multiple devices attempt to transmit simultaneously.
802.11e: EDCF

• **ACs are differentiated based on their CSMA parameters:**
  - **IFS**
  - **CWmin**
  - **CWmax**
  - **Backoff exponent**
802.11e: EDCF

- Higher priority ACs are assigned parameters that result in shorter CWs so that a statistical advantage is gained in accessing the channel.

- Protocol parameters become vectors:
  - $CW_{\text{min}}[AC]
  - $CW_{\text{max}}[AC]
  - $\text{AIFS}[AC]
  - $bck[AC]

- $CW[AC,t]$ is derived with the usual CSMA/CA rules.
802.11e: EDCF

- Arbitration InterFrame Space (AIFS) substitute the common DIFS
- Each AIFS is at least DIFS long

- Before entering the backoff procedure each Virtual Station will have to wait AIFS[AC], instead of DIFS
Arbitration IFS (AIFS)

802.11a: slot=9 μs, SIFS=6 μs, PIFS=15 μs, DIFS=24 μs, AIFS ≥34 μs
Contestation Window

- $CW_{\text{min}}[AC]$ and $CW_{\text{max}}[AC]$

- Contention Window update:

\[
CW_{\text{new}}[AC] = (CW_{\text{old}}[AC] + 1) \cdot bck - 1
\]
Backoff

802.11a: slot=9 $\mu$s, SIFS=16 $\mu$s, PIFS=25 $\mu$s, DIFS=34 $\mu$s, AIFS $\geq$34 $\mu$s
Virtual Stations

- Each AC queue behaves like a different virtual station (independent sensing and backoff)

- If the backoff counters of two or more parallel ACs in the same QSTA reach 0 at the same time, a scheduler inside the QSTA avoids virtual collision by granting the TXOP to the AC with the highest UP

- The lowest priority colliding behaves as if there were an external collision
802.11e: EDCF – Beacon Frames

• Values of \( AIFS[AC] \), \( CW_{\text{min}}[AC] \) e \( CW_{\text{max}}[AC] \) are determined by the QAP and transmitted within beacon frames (normally every 100 msec)
• QSTAs must abide to the received parameters
• QSTAs may use these parameters to chose the QAP the prefer to connect to (estimate of the expected performance)
802.11e: EDCF – Virtual Collisions

• Every AC within a QSTA behaves as if it were an independent station, with its own MAC parameters AIFS[AC] e CW[AC]
• So Virtual Stations (AC queues) within a QSTA contend for the channel
• Internal collisions between different ACs are solved virtually, without loss of resources
• The TXOP goes to the highest priority AC and the others behave as if there was a real collision
802.11e: EDCF – Virtual Collisions

- **AIFS[TC]**
- **AIFS[TC] = DIFS**
- **PIFS**
- **SIFS**
- **ACK**
- **DATA**
- **high priority TC**
- **medium priority TC**
- **low priority TC**
- **RTS**
- **CTS**
- **Contention Window** (counted in slots, 9us)
- **defer access**
- **count down as long as medium is idle, backoff when medium gets busy again**

With 802.11a:
- Slot: 9us
- SIFS: 16us
- PIFS: 25us
- DIFS: 34us
- AIFS: >=34us
802.11e: TXOP

- TXOP is the time interval in which a STA may use the channel.
- It’s an initial time plus a duration, indeed the negotiation is no more for a PDU, but can be for many aggregated PDUs (this part is not well defined in the standard).
- $CW[AC]$ is managed with usual rules of increment (after collisions/failures) and decrement (during idle channel):

$$NewCW[AC] = ((OldCW[AC] + 1) \times 2) - 1$$
802.11e: EDCF

- Sample allocation of TCId to ACs:

<table>
<thead>
<tr>
<th>TCID</th>
<th>CA</th>
<th>Traffic description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>Best Effort</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>Best Effort</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>Best Effort</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>Video Probe</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>Video</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>Video</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>Voice</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>Voice</td>
</tr>
</tbody>
</table>
EDCA Bursting

- Once the station has gained access to the medium, it can be allowed to send more than one frame without contending again.

- The station cannot transmit longer than TXOP_Limit.

- ACK frame by frame or Burst ACK.

- SIFS is used between packets (to avoid collisions).
EDCA Bursting: Pros / Cons

• **Pros**
  - Reduces network overhead
  - Increases throughput (SIFS and burst ACKs)
  - Better fairness among the same priority queues: independently of the frame size, a QSTA gets a TXOP every time it wins a contention
    • E.g., STA A uses 500 B frame; STA B uses 1K B frame. Thus B would get higher throughput in 802.11, while in 802.11e both can get approximately same throughput
EDCA Bursting: Pros / Cons

• **Cons**
  
  - Possible increasing of delay jitter
  
  - TXOP\_Limit should not be longer than the time required for transmitting the largest data frame

• In any case EDCA does not solve the downlink/uplink unfairness problem
802.11e: HCF

- HC may allocate TXOPs to himself (QAP) or to other QSTAs
- Self allocation is done to transmit MSDUs, allocation of resources may solve the uplink/downlink unfairness
- Allocation to AP can be done after a Point coordination InterFrame Space (PIFS) con PIFS < DIFS
- HC (QAP) has priority over other stations and may interrupt a CP to start a CFP transmitting a Poll frame
802.11e: HCF

• Time is divided between contention free periods (CFP) and contention periods (CP), that are alternated roughly cyclically.

• A sequence CFP + CP defines a Periodic Superframe of 802.11e.

• The CP can be interrupted by other contention free periods called CAPs.
802.11e: HCF
MAC 802.11e: HCCA

![Diagram of MAC 802.11e: HCCA]
802.11e: HCF – QoS CFPoll Frame

- Within a CP, TXOP is determined either:
  - Through EDCF rules (free channel + AIFS + BO + TXtime)
  - Through a poll frame, called QoS CFPoll, sent by HC to a station
- QoS CFPoll is sent after PIFS, so with priority wrt any other traffic
- Indeed there is not a big difference between a CFP and CAPs as defined above.
802.11e: HCF – QoS CFPoll Frame

- During CFP, TXOPs are again determined by HC and QoS CFPoll can be piggybacked with data and ACKs if needed
- Stations not polled set NAV and cannot access the channel
- The CFP must terminate within a time specified within the beacons and it is terminated by the CF-End frame sent by HC
802.11e: HCF – QoS CFPoll Frame

- QoS CF-Poll frame was introduced with the 802.11e amendment, for backward compatibility it contains a NAV field the legacy stations can use to avoid interfering.
- NAV specify the whole TXOP duration.
- Legacy stations in HCF can only use the CP period.
802.11e: HCF – Controlled Content.

• Is a mix between polling and contention based
• Should guarantee better e-performances than contention during congestions periods
• The *Controlled Contention* mechanism is mandatory for QAP an optional for QSTA:
  - QSTA notify QAP some allocation requests, QAP will allocate the necessary TXOPs via polling
  - Different from standard polling, because it’
802.11e: HCF – Controlled Content.

- QAP defines if there are resources to satisfy requests:
  - If available schedules the channel (IEEE 802.11e does not specify scheduling algorithms, these are open for research and competitive implementation)
  - The answer to stations can be acceptance, rejections or a proposal to use resources with a lower priority
HCCA

- HCCA effectively provides policing and deterministic channel access by controlling the channel through the HC
- It is backward compatible with basic DCF/PCF
- Based on polling of QSTAs by the HC
Crucial features of HCCA

- HCCA operates in CP and CFP
- Uses TXOPs which are granted through HC (in HCCA!)
  - HC allocates TXOPs by using QoS CF-Poll frames
  - In CPs, the time interval during which TXOPs are polled by HC is called CAP (Controlled Access Period)
  - 8 Traffic Categories (TCs)
HC Behavior in HCCA

- According to HCCA:
  - HC may allocate TXOPs to itself to transmit MSDUs whenever it wants, however only after having sensed the channel idle for PIFS
  - In CP, the HC can send the CF-Poll frame after a PIFS idle period, thus starting a CAP
  - In CFP, only the HC can grant TXOPs to QSTAs by sending the CF-Poll frame
  - The CFP ends after the time announced by HC in the beacon frame or by the CF-End frame from HC
QSTA Behavior in HCCA

• A QSTA behaves as follows
  - In CP QSTAs can gain a TXOP thanks to a CF-Poll frame issued by HC during CAPs, otherwise they can use EDCA
  - In CFP, QSTAs do not attempt accessing the channel on their own but wait for a CF-Poll frame from the HC

• The HC indicates the TXOP duration to be used in the CF-Poll frame (QoS-control field)
  - Legacy stations kept silent by NAV whenever they detect a CF-Poll frame
During the CP, a TXOP may begin because:

- The medium is determined to be available under EDCA rules (EDCA-TXOP)
- The STA receives a special polling frame from HC (polled-TXOP)
Polling in HCCA

- Polling list is a crucial key in HCCA
  - Traffic scheduling (i.e., how QSTAs are polled) is not specified
  - QSTAs can send updates to the HC on their queue size as well as on the desired TXOP, (through the QoS control field in data frames)
  - QSTAs can send ADDTS requests to initiate a new traffic stream
Traffic Signaling

- Two types of signaling traffic are supported:
  - Connectionless queue state indicator
    - E.g., Arrival rate measurement: notification and not negotiation between peer entities is used
  - TSPEC (Traffic Specification) between HC and QSTAs
    - E.g., service negotiation and resource reservation
Traffic Signaling

- TSPEC are the base for CAC
- QoS without CAC is impossible
- QoS is granted to flows not to packets
- Flows are persistent (normally)
- Flows can be predicted (sometimes)
Resource Scheduling

- Not essential to backward compatibility
  - The standard has just a reference impl. (SS)
- HCF is implemented in the AP
  - HCCA scheduling is a function of HCF
- Requirements of traffic flows are contained in the Traffic Specifications (TSPEC):
  - Maximum, minimum and mean datarate
  - Maximum and nominal size of the MSDUs
  - Maximum Service Interval and Delay Bound
  - Inactivity Interval
  - ...

Renato.LoCigno@dit.unitn.it
Thanks: TLC Networks Group – Politecnico di Torino

Nomadic Communications: 802.11e 51
EDCA Differentiation
HCCF Scheduling

Renato Lo Cigno
www.disi.unitn.it/locigno/didattica/NC/
Quest'opera è protetta dalla licenza Creative Commons NoDerivs-NonCommercial. Per vedere una copia di questa licenza, consultare: http://creativecommons.org/licenses/nd-nc/1.0/ oppure inviare una lettera a:
Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.

This work is licensed under the Creative Commons NoDerivs-NonCommercial License. To view a copy of this license, visit: http://creativecommons.org/licenses/nd-nc/1.0/ or send a letter to Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.
Thanks & Disclaimer

• These slides and results are based on the following paper

• As such they must be considered examples of the possible performances and tradeoffs.
• Thanks to Bianchi and all the other authors for providing copy of the papers, graphics and slides.
EDCA or HCCA?

• How does EDCA support differentiation?
• Is this enough for standard purposes?
• Are parameters easy to tune and universal?

• How can HCCA polling-based scheduling be implemented?
• Do we need to use the feedback from the STA?
• How can the traffic be described?
Performance Evaluation of Differentiated Access Mechanisms Effectiveness in 802.11 Networks

G. Bianchi, I. Tinnirello, L. Scalia

presented @ Globecom 2004
QoS Support issues in legacy 802.11

• DCF is long term fair
  • Equal channel access probability among the stations
  • Averagely, the same channel holding time (for homogeneous packet sizes)
    • Solution: differentiate packet sizes?
    • Solution: differentiate channel holding times?

• NO WAY! QoS is not a matter of how long I hold the channel
  • It means more...
    • Need to manage access delay problems for real-time apps!!!
    • Need to modify 802.11 channel access fairness!!!
QoS @ IEEE 802.11 MAC

- 802.11e defines different traffic classes onto map data flows
- Each traffic class behaves as an independent MAC entity
- Differentiated access priority is provided by:
  - Giving probabilistically lower backoff counters (\(CW_{\text{min}}, CW_{\text{max}}, PF\))
  - Giving deterministically lower inter-frame spaces and backoff de-freezing times. (\(\text{AIFS}_N\))

EDCA

Different MAC Access Parameters @ each class to differentiate channel access probability

Backoff based parameters: \(CW_{\text{min}}, CW_{\text{max}}, PF\)
Channel monitoring based parameters: \(\text{AIFS}\)
EDCA Performance Evaluation

- Performance Evaluation: answers we try to give...
  - Homogeneous sources
    - Performance effectiveness of each differentiation MAC parameter, individually taken
    - How each differentiation parameter reacts to different load conditions?
  - Heterogenous sources
    - What are the most effective settings to manage high-priority delay requirements?
EDCA Performance Evaluation

• Simulations
  – Same number of HP and LP stations
  – Same packet size (1024 bytes)
• Homogeneous sources scenario
  – Saturation conditions for HP and LP stations
    • Queues never empty
    • Data rate = Phy rate = 1 Mbps
• Heterogeneous sources scenario
  – 3 pkts/sec. for HP traffic
  – Saturation conditions for LP traffic
    • Data rate = Phy rate = 1 Mbps
**CWmax Differentiation (1)**

- **CWmax(A) < CWmax(B)**
  - Once reached **CWmax** (repeated collisions), A gets access priority over B

---

A extracts probabilistically a lower backoff value due to its lower **CWmax**.
CWmax Differentiation (2)
**CWmax Differentiation (3)**

- **Low throughput differentiation**
  - Only with CWmax=64 effective
  - @ low loads poor performance
  - Few collisions
- **Inefficient channel usage**
  - Consecutive Collisions are needed for the differentiation effect
  - Overall throughput suffers @ high loads
PF Differentiation (1)

- PF(A) < PF(B)
  - once a collision occurs, station A has probabilistically an
    higher chance to extract a lower backoff value, thus it may
    retransmits first.

A extracts probablistically a lower backoff value due to its lower CW
PF Differentiation (2)
PF Differentiation (3)

- PF is greater than 2 for LP stations.
- $\text{CW}_{\text{new}} = \text{PF} \times \text{CW}_{\text{old}}$
- It is sufficient a single collision to begin the differentiation process.
- Impossible to force LP traffic to zero!
  - After a packet successful transmission, the PF effect is no more present
CWmin Differentiation (1)

- CWmin(A) < CWmin(B)
  - In average, station A has a lower backoff than B

Thanks to its lower CWmin, A extracts probabilistically a lower backoff value.
CWmin Differentiation (2)
CWmin Differentiation (3)

- Very High differentiation performance
- @ low loads performance is good
  - Collision effects among HPs not significant
- @ high loads collisions mainly involve HP stations (because of their small CW)
  - Degradations regard HP traffic -> bad!
  - LP traffic not affected
- Collision effects un-altered
AIFS Differentiation (1)

- AIFS(A) < AIFS(B)
  - station A decrements its backoff timer before than station B

Thanks to its lower AIFS, A starts decrementing its backoff value before than B either after busy channel or idle channel conditions.
AIFS Differentiation (2)
AIFS Differentiation (3)

- Very High differentiation performance
  - Complementary to CWmin case
- @ low loads differentiation performance suffers
  - Collision are few ->
- @ high loads collisions mainly involve LP stations, since HP stations access first
  - Degradations regard LP traffic -> good!
  - HP traffic not affected
Heterogeneous Sources: Throughput

- Focus on AIFS and CWmin differentiation, seen to be most effective

The minimum differentiation effect allows to guarantee HP traffic!!!
1) CWmin more effective to manage delay behaviour than AIFS (see slopes)
2) AIFS differentiation slightly sensitive to load in terms of delay
3) Joint use: delay requirements satisfied with AIFS, throughput managed via CWmin (because of the maxima)
Conclusions

• Cwmin and AIFS differentiation perform better than PF and CWmax differentiation
  – PF and CWmax differentiation operations allowed only by collisions
• CWmin and AIFS show a complementary behaviour
  – CWmin performance degrades @ high loads
  – AIFS performance degrades @ low loads
• Joint use of CWmin and AIFS
  – AIFS to meet delay requirements
  – CWmin to manage throughput performance
• Complex parameter setting
• Behavior hardly predictable