Wireless LANs and QoS

Prof. Andrzej Duda
duda@imag.fr

http://duda.imag.fr

Global overview

- Recall on 802.11
- QoS in WLANs
- Performance of 802.11 DCF
  - 802.11 performance anomaly
  - Scheduling time-sensitive traffic on 802.11
  - Short-term fairness of 802.11
- 802.11 PCF
- 802.11e
- Other proposals and evaluation
- Conclusion

2

802.11

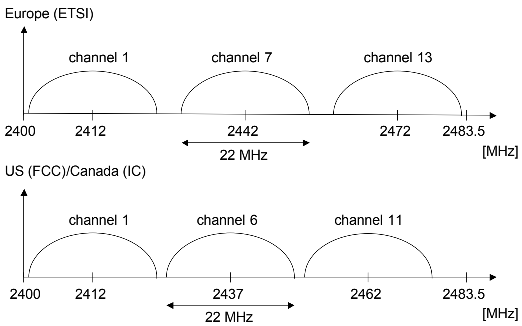
3

802.11 - Physical layer

- 802.11b
  - frequency band of 2.4 GHz: [2,4 GHz ; 2,48 GHz]
  - nominal bit rate of 11 Mb/s
  - passes through concrete
- 802.11g
  - frequency band of 2.4 GHz
  - nominal bit rate of > 22 Mb/s
- 802.11a
  - frequency band of 5 GHz: [5,15 GHz ; 5,825 GHz]
  - nominal bit rate of 54 Mb/s
    - 6, 9, 12, 18, 24, 36, 48, 54 Mb/s, (6, 12, 24 Mb/s mandatory)
  - LOS - Line-of-Sight (no obstacles)

4

Channel selection



5

802.11 - MAC layer

- Same MAC layer for all variants (802.11b, a, g)
  - DCF (Distributed Coordination Function)
    - access method - CSMA/CA (Carrier Sense Multiple Access/Collision Avoidance)
    - similar to Ethernet, no collision detection
  - PCF (Point Coordination Function)
    - polling, optional

6

802.11b

- 802.11b: wireless LAN
  - nominal bit rate of 11 Mb/s, degraded to 5.5, 2, 1 Mb/s
  - shared radio channel
- MAC layer
  - DCF (Distributed Coordination Function)
    - access method - CSMA/CA (Carrier Sense Multiple Access/Collision Avoidance)
    - similar to Ethernet, no collision detection
  - PCF (Point Coordination Function)
    - polling, optional

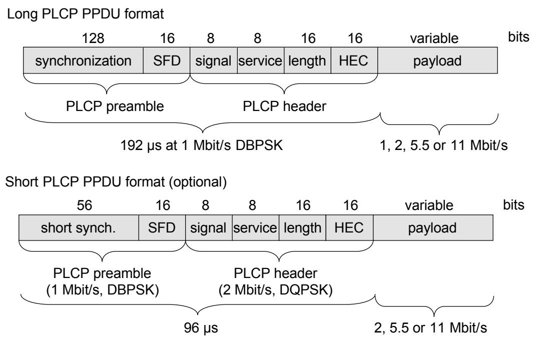
7

Rates of 802.11b

Rate	Code	Modulation	Symbol rate	Bits per symbol
1 Mbps	11 bits Barker	BPSK	1 MSps	1
2 Mbps	11 bits Barker	QPSK	1 MSps	2
5.5 Mbps	8 bits CCK	QPSK	1.375 MSps	4
11 Mbps	8 bits CCK	QPSK	1.375 MSps	8

8

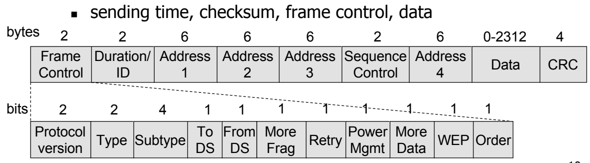
802.11b - Physical layer frames



9

802.11 - Frame format

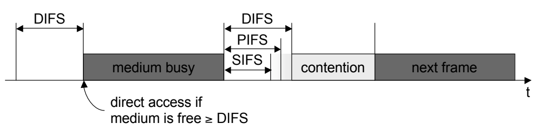
- Types
  - control frames, management frames, data frames
- Sequence numbers
  - important against duplicated frames due to lost ACKs
- Addresses
  - receiver, transmitter (physical), BSS identifier, sender (logical)
- Miscellaneous
  - sending time, checksum, frame control, data



10

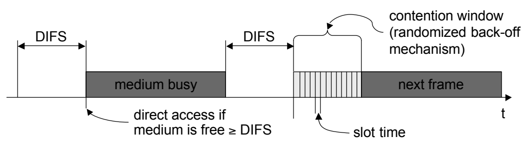
802.11 MAC

- Inter-frame spacing
  - SIFS (Short Inter Frame Spacing)
    - for ACK, CTS, polling response
  - PIFS (PCF IFS)
    - for time-bounded service using PCF
  - DIFS (DCF IFS)
    - for contention access



11

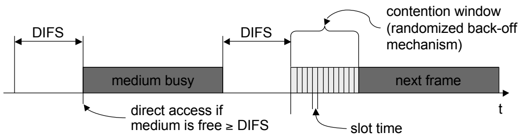
802.11 DCF - CSMA/CA



- Channel idle during DIFS, transmit frame
- If the medium is busy, wait for a free DIFS and a random back-off time (collision avoidance, multiple of slot-time)
- If another station uses the medium during the back-off time of the station, the back-off timer stops (fairness)

12

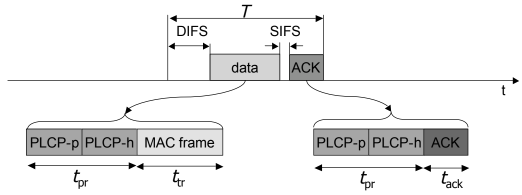
802.11b DCF - CSMA/CA



- Back-off time - random interval
  - Contention Window - uniform distribution [0, CW] * SLOT
  - doubled if collision - CW: CW_{min} = 31, CW_{max} = 1023
  - SLOT = 20 μs
- If another station uses the medium during the back-off time of the station, the back-off timer stops (fairness)

13

802.11b - transmission



- *DIFS* = 50 μs, *SIFS* = 10 μs
- t_{ack}= 10 μs, if 11 Mb/s and ACK frame size 112 bits
- PLCP (Physical Layer Convergence Protocol) preamble and header: t_{pr}= 192 μs, if 1 Mb/s, t_{pr}= 96 μs, if 11 Mb/s
- 270 μs for any payload, time to transmit 370 B at 11 Mb/s

14

Single host performance

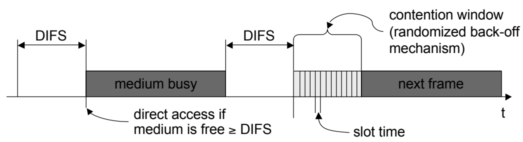
- Useful throughput proportion

$$t_{tr}/T$$
$$T = t_{tr} + t_{ov}$$
$$t_{ov} = DIFS + t_{pr} + SIFS + t_{pr} + t_{ack}$$

- Useful throughput proportion if 1500 B of data
  - 0.70
  - throughput of 7.74 Mb/s
  - measured at UDP layer 5.08 Mb/s

15

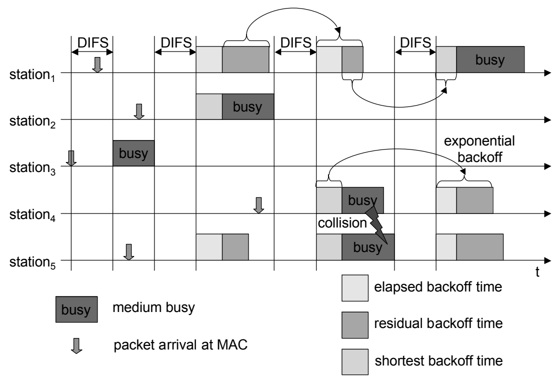
802.11g DCF - CSMA/CA



- Back-off time - random interval
  - Contention Window - uniform distribution [0, CW] * SLOT
  - doubled if collision - CW: CW_{min} = 15, CW_{max} = 1023
  - SLOT = 9 μs, if rate > 11 Mb/s

16

802.11 - contention



17

QoS in WLANs

18

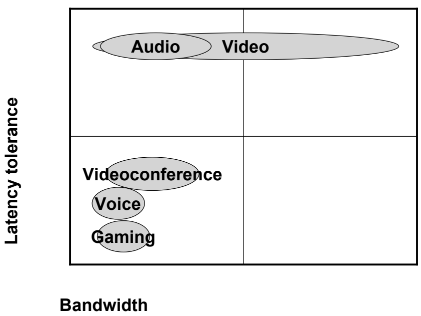
What does QoS mean?

- We limit the definition to mean "delivering traffic for real-time applications"
- Define a set of applications first
  - voice
  - gaming
  - real-time video (videoconferencing)
  - "CD like" audio
  - "Television/VCR like" video
- Each application has a requirements tuple
  - max latency
  - min data rate
  - max packet drop probability

Application requirements

- Video:
  - mean rate 2Mb/s (SD) to 20Mb/s (HD), variable
- Audio
  - bandwidth range 64kb/s up to 1.5Mb/s
  - relatively high latency tolerance
- Voice and videoconferencing
  - requires low latency (< 50ms), but can tolerate frame losses
  - 32kb/s and lower for voice, 128kb/s for videoconferencing
- Gaming
  - lowest latency (< 10ms) required
  - low tolerance to frame loss
  - lower bandwidth requirements (32kb/s – 128kb/s)

Applications summary

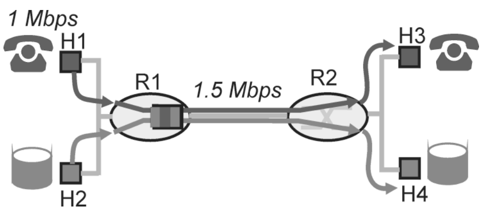


QoS principles

- Marking of frames is needed to distinguish between different QoS classes
  - 802.1Q tags, 802.1D forwarding operation
- Scheduling of different classes
  - Priority Queue, Weighted Round Robin, Weighted Fair Queue
  - need for traffic isolation
- Resource allocation
  - what part of resources is given to each class?
- Admission control
  - admit flows to fit available resources (how to reject others?)
  - require policing mechanisms to ensure the sources adhere to their allocations (how to drop excess traffic?)
- Alternative to statistical multiplexing
  - fixed reservation and allocation (802.16 TDMA on demand)

QoS principles

- Marking
- Scheduling of different classes and resource allocation
- Admission control



802.1D Traffic Classes

- **Network control** (7): time-critical and safety-critical
- **Voice** (6): Time-critical, less than 10 ms delay
- **Video** (5): Time-critical, characterized by less than 100 ms delay, such as interactive video
- **Controlled load** (4): loss-sensitive, such as streaming multimedia and business-critical traffic
- **Excellent effort** (3): loss-sensitive, but of lower priority than controlled load
- **Best effort** (2): Not time-critical or loss-sensitive
- **Background** (0): Not time-critical or loss-sensitive, and of lower priority than best effort

WLAN QoS

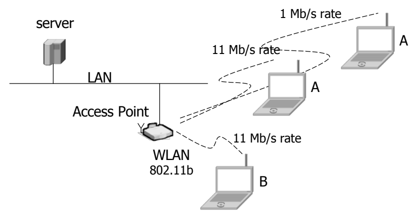
- Wireless is different than wired!
  - "rubber pipe problem" – channel conditions vary resulting in
    - varying rates
    - temporarily very high error rates (10 – 20%)
- Wireless is unreliable!
  - irrespective of channel access scheme or protocol, each transmission attempt may fail
  - failures are detectable, and data can be retransmitted
  - if the application can wait, a packet may be transmitted multiple times to increase probability of success
- Conclusion
  - There is no such thing as "guaranteed QoS" in a WLAN. All QoS schemes in a WLAN are statistical. Some schemes have better statistics than others
  - What does "connection admission control" mean in an unlicensed RF band?

25

Performance anomaly of 802.11

26

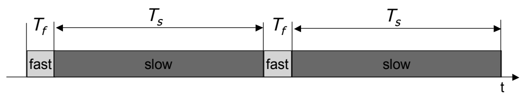
Motivation



nominal rate (Mb/s)		throughput (Mb/s)	
A	B	A	B
11	11	3.09	3.36
1	11	0.76	0.76

27

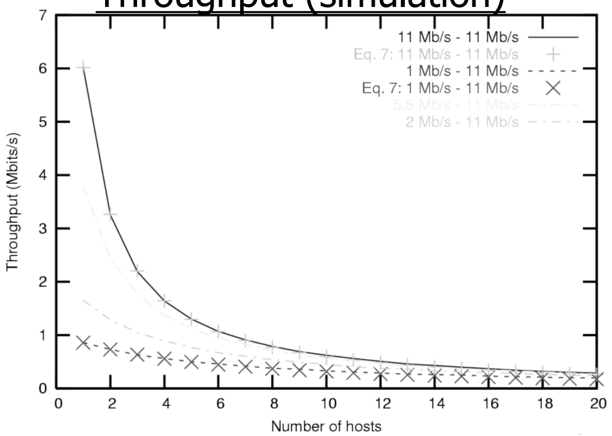
One slow host, one fast



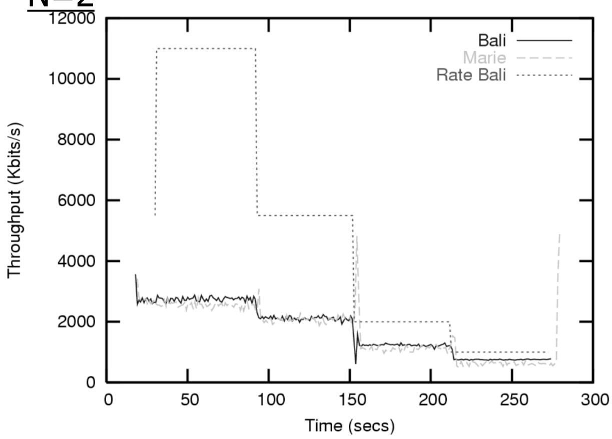
- CSMA/CA
  - equal access probability, not equal share
- Throughput when R = 11 Mb/s, r = 1 Mb/s
  - Fast: (1/11)/(1/11 + 1) * 0.7 * 11 Mb/s = 0.64 Mb/s
  - Slow: (1)/(1/11 + 1) * 0.7 * 1 Mb/s = 0.64 Mb/s

28

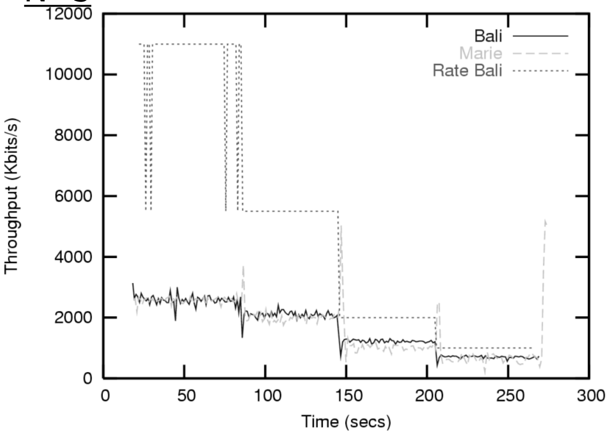
Throughput (simulation)



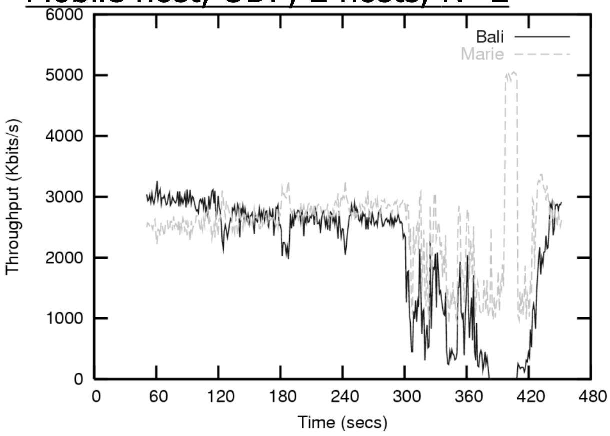
N=2



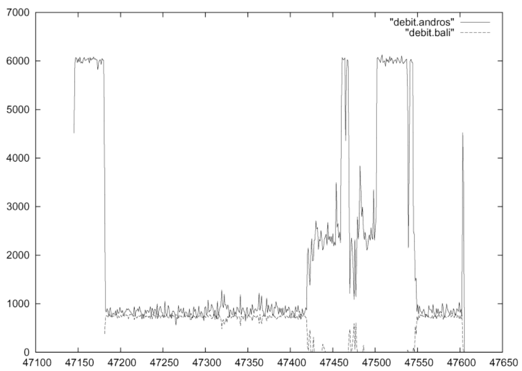
N=3



Mobile host, UDP, 2 hosts, N=2

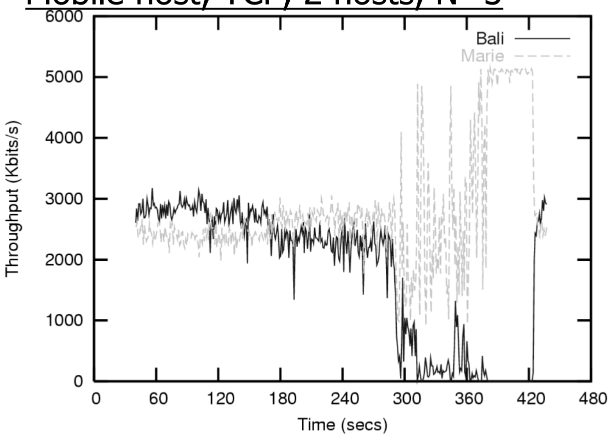


Mobile host, UDP, hosts at 11 Mb/s and 1 Mb/s moving, N=2



33

Mobile host, TCP, 2 hosts, N=3



Performance of 802.11

- Analysis of the overhead of 802.11b DCF
  - important overhead, useful 70%
- Greedy hosts with different rates
  - analytical explanation of the phenomenon
- Comparison with measurements
  - enforced lower rate and real mobility
- Strong influence of a slow host on overall performance
- Other factors alleviate the problem
  - TCP congestion control

35

Scheduling time-sensitive traffic on 802.11

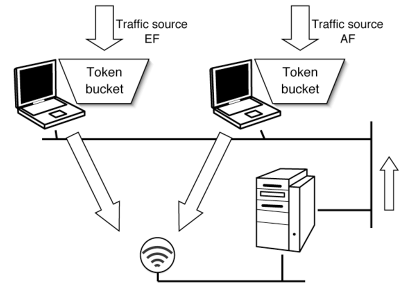
36

Scheduling time-sensitive traffic on 802.11

- Common wisdom
  - one cannot guarantee small delays on 802.11 DCF
- Two classes (*DiffServ* like), two hosts
  - time-sensitive EF class, higher priority
  - greedy AF class, lower priority
- Experiments with constrained sources
  - keep channel non saturated
    - time-sensitive EF sources can obtain small delays even if competing with greedy AF sources
  - if saturated channel
    - still possible for time-sensitive EF sources to obtain small delays provided their packet rate remains lower than a limit packet rate

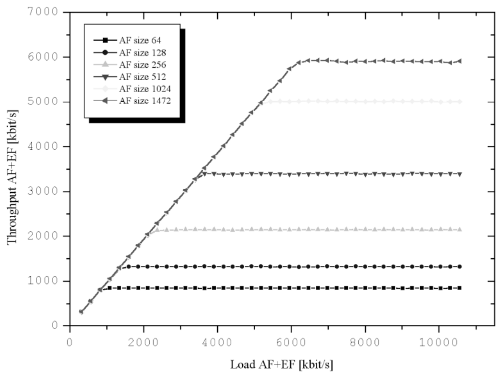
37

Experimental set up



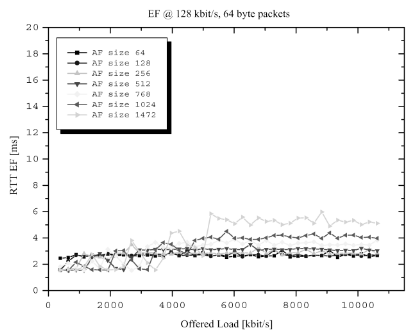
38

Constant 64 kb/s EF traffic



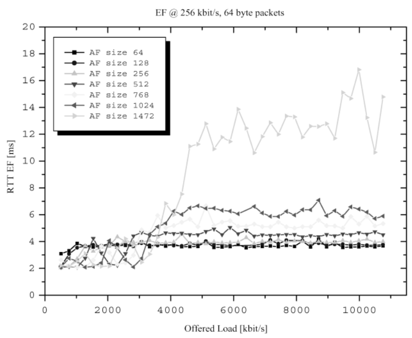
39

Small EF rate



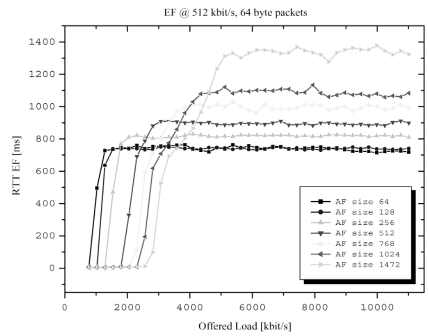
40

Medium EF rate



41

High EF rate



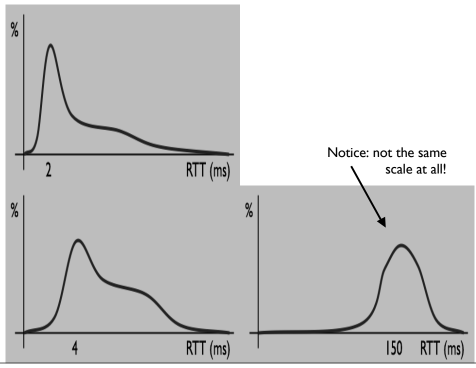
42

EF RTT: 3 cases

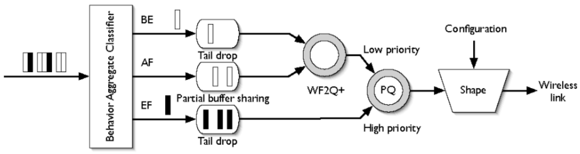
before saturation by AF

after saturation by AF

Notice: not the same scale at all!



DiffServ based QoS



- *DiffServ* scheduling in hosts
  - if EF packet rate < x^{sat} then EF will have low delays even in saturation
  - if EF packet rate > x^{sat} then limit the AF traffic, so EF can obtain low delays

44

Wrap-up

- QoS scheduling in a 802.11 WLAN
  - learn EF and AF characteristics (packet rates and sizes)
  - decide if EF can be accepted, or
  - limit AF traffic, so that EF never becomes saturated
- Interesting for adaptive sources
  - change codec, if the limiting rate decreases
- Results of measurements for 2 hosts
  - for N > 2, collision probability increases
  - some hosts double the congestion window - the delay distribution will be shifted towards higher delays

45

Short-Term Fairness of 802.11

46

Fairness of 802.11

- Paper by Koksal *et al.* 2000 (MIT)
  - study of WaveLAN cards, fairness index, strong short term unfairness
- Ethernet capture effect
  - station that transmits after backoff has more chances to transmit once again (CSMA/CD)
- Common wisdom
  - 802.11 CSMA/CA is fair in a long term
  - it presents short term unfairness
- Not true!
  - compared with Slotted ALOHA, 802.11 presents even better (N = 2) or similar fairness (N > 2)

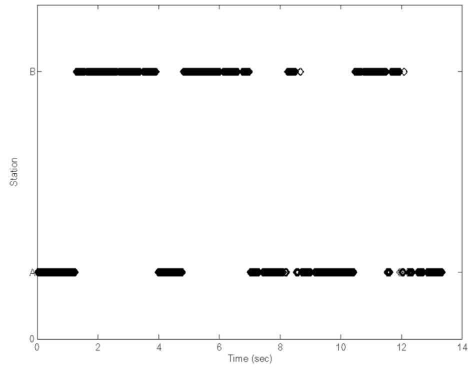
47

Fairness

- Desirable property of a MAC layer
- Provide equal channel shares among stations
- Long-term ≠ Short-term fairness
  - Short-term ⇒ long-term fairness
  - Short-term fairness ⇒ low latency

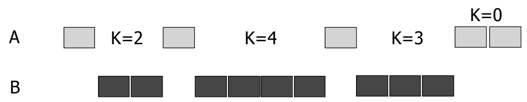
48

Wavelan channel capture (Koksal)



49

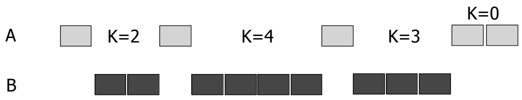
Inter-transmission distribution



- K - number of transmissions by other hosts between two successful transmissions of a given host
- Two hosts A and B, observe A
  - K=0 - next transmission done by A,
  - K=1 - B transmits once, next transmission done by A,
  - K=2 - B transmits twice, next transmission done by A, and so on.

50

Inter-transmission distribution



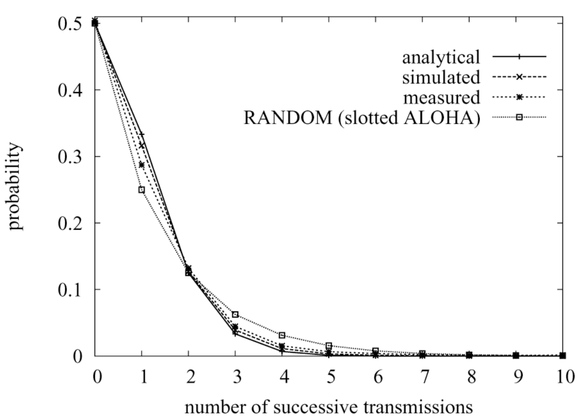
- Large values of K mean lower short-term fairness, because a host may capture the channel for several successive transmissions
- For N hosts
  - **mean**: E(K) = N - 1, good fairness (N-1 for TDMA)
  - E(K) > N-1 ⇒ poor (long-term) fairness
  - E(K) < N-1 ⇒ short tail, good fairness
  - **capture probability**: P(0) = 1/N, equal access probability

51

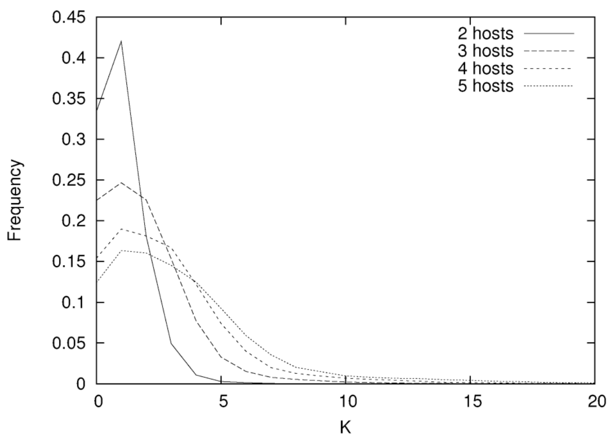
Inter-transmission distribution

- TDMA, 2 hosts
  - P(K=k) = 0, if k=0, 1 if k=1
  - E(K) = 1
- RANDOM (Slotted ALOHA), 2 hosts
  - P(K=k) = $1/2^{k+1}$, k=0, 1, …
  - P(0) = 1/2
  - E(K) = 1
- 802.11 DCF (2 hosts), approximated (no collisions)
  - P(K=k) = (k+1)/(k+2)! , k=0, 1, …
  - P(0) = 1/2
  - E(K) = 2 - e = 0.718

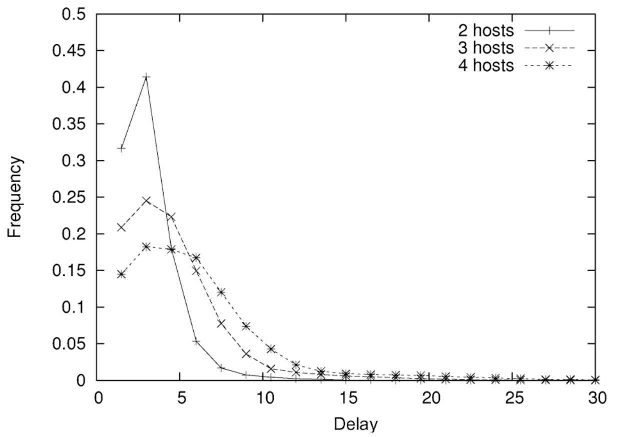
52



53



54



55

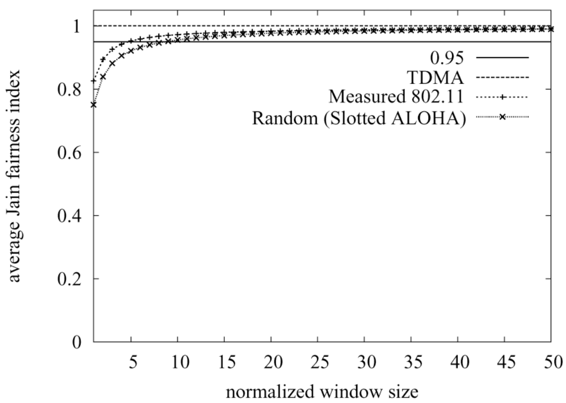
Jain fairness index

$$F_J(w) = \frac{(\sum_{i=1}^{N} \gamma_i)^2}{N \sum_{i=1}^{N} \gamma_i^2}.$$

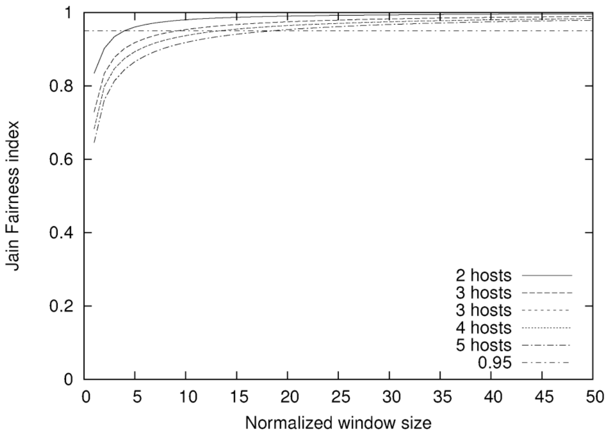
- AAAB AAAA BBAA
  - $w = 4$
  - $\gamma_A = 0.75$, $\gamma_B = 0.25$
  - $\gamma_A = 1$, $\gamma_B = 0$
  - $\gamma_A = 0.5$, $\gamma_B = 0.5$

- Jain fairness index
  - γ_i - proportion of channel allocations over time intervals
  - $F_J(w) = 1$ perfect fairness
  - $F_J(w) = 1/N$ perfect unfairness

56



57



58

Short-Term Fairness of 802.11

- Short term fairness is good for small number of hosts
  - N = 2, better than Slotted ALOHA
- For larger number of hosts, short term fairness slightly degrades
  - some hosts collide and double their contention window
  - still better that Wavelan analyzed in the Koksal paper
- Short term fairness essential for obtaining low delays
- Common wisdom comes from the confusion of the access methods
  - Wavelan - doubles contention window when channel busy
  - 802.11 - doubles contention window when collision

59

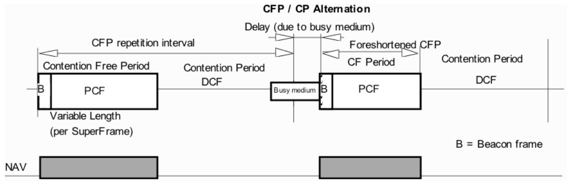
Performance of 802.11

- 802.11 presents several good performance characteristics
  - good short term and long term fairness for small number of hosts (802.11b optimal for N=3, 4)
  - the residual congestion window improves short-term fairness
  - behaves like a probabilistic Round-Robin providing equal packet rates to hosts
  - time-sensitive flows can obtain low delays even in saturation
- Drawback
  - "unfair" sharing when different rates
  - becomes even more critical for 802.11a and 802.11g (rates from 6 Mb/s to 54 Mb/s)

60

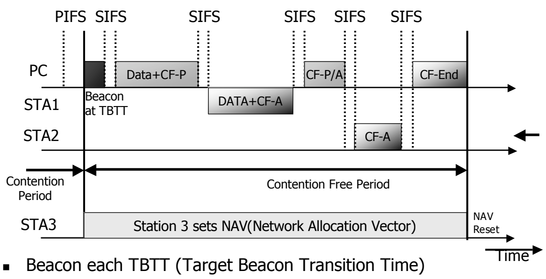
802.11 PCF

PCF (Point Coordination Function)



- Single Point Coordinator (PC) controls access to the medium (Access Point acts as PC)
- PC transmits beacon packet when medium is free for time greater than PIFS (higher priority than the DCF)
  - beacon contains length of CFP
- During PCF mode
  - PC polls a station in a list for data
  - after a transmission of a MPDU, move on to the next station

PCF



- Beacon each TBTT (Target Beacon Transition Time)
- PC polls stations or ends with CF-End
- POLL/ACK piggybacking
  - Data+CF-Poll, Data+CF-ACK+CF-Poll
  - CF-ACK (nodata), CF-Poll (nodata), CF-ACK+CF-Poll

Limitations of PCF

- Unpredictible beacon delay
  - on-going transmissions need to complete (4.9 ms in the worst case 802.11a)
- Access may not be exactly periodic due to polled users
  - entering/exiting polling list
  - transmitting data of varying packet sizes - maximal frame length (2312 Bytes) at a chosen bit rate
- PCF may be forced to end early
  - PC do not serve some members of the polling list (must wait)
- Ineffective on large networks
  - long list of polled stations
- Hidden station problem
  - misses the previous beacon and operates as in DCF
  - can transmit interfering frames during CFP

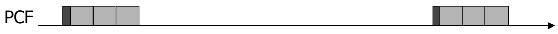
802.11e

QoS strategy

- Use 802.1Q/802.1D tags to classify traffic into groups with widely differing requirements
- 8 priority levels (TC) grouped into four classes
  - best effort
  - video/audio probe
  - video/audio
  - voice/gaming
- Four parameters to vary (EDCF):
  - CWmin/CWmax (weak)
  - AIFS (Arbitration Inter-Frame Space) (strong)
  - TXOP duration (strong)
- Each of four classes gets a different set of parameters, specified by the AP or set to defaults
- Coordinator polls for transmission (HCF)

802.11e hybrid strategy

- most efficient, low latency for bursty traffic; per packet overhead

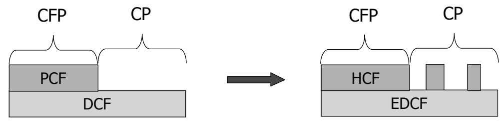


- efficient for predictable traffic, inefficient for bursty traffic

- short PCF-type bursts provide efficient access for predictable traffic, DCF-type operation provide efficient access for bursty traffic

67

802.11e

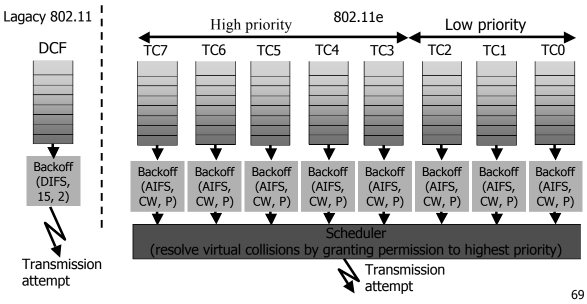


- Two coordination functions:
  - EDCF (Enhanced DCF), during contention period
  - HCF (Hybrid Coordination Function), during contention period and contention free period

68

EDCF

- Eight Traffic Categories (TCs) within one station
- for 802.11a: CWmin(TC) = 0-255, P(TC) = 1-16



69

EDCF

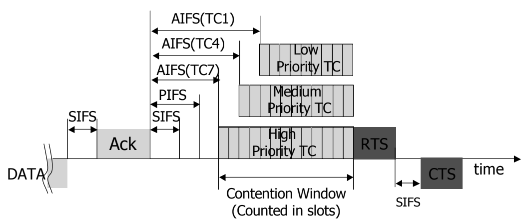
- TXOP (Transmission Opportunity)
  - an interval of time when a particular STA has the right to initiate transmissions
    - defined by a starting time and a maximum duration
  - allocated via contention (EDCF-TXOP) or granted through HCF (polled-TXOP)
  - a station may transmit multiple MAC frames (with a SIFS gap between ACK) in a TXOP
- EDCF-TXOP
  - wait for AIFS(TC)
  - random number of slots from [1, CW+1]
  - CWmin(TC)
  - persistence factor P(TC)

new CW = [(old CW+1)*P(TC)] - 1

  - P = 2 for 802.11

70

IFS in EDCF



- For 802.11a:
  - slot 9μs
  - SIFS 16 μs, PIFS 25 μs, DIFS 34 μs, AIFS ≥ 34 μs,

71

EDCF characteristics

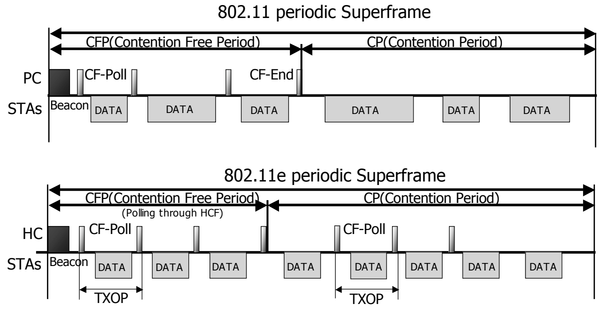
- Priority access improves chances of getting access to the medium quickly
- Long burst duration provides high bandwidth access, but at the expense of latency
- Set appropriately:
  - voice/gaming has very high access priority, small burst size
  - video/audio has much lower access latency (but better than best effort), but large burst sizes

72

HCF

- During CFP
  - Poll STAs and give a station the permission to access channel
  - Starting time and maximum duration of each TXOP are specified by HC
- During CP
  - HC can issue polled TXOPs in the CP by sending CF-Poll after a PIFS idle period

802.11e superframe

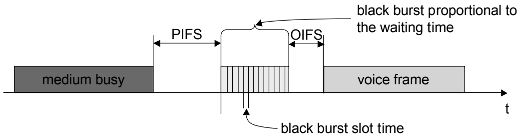


Traffic signaling

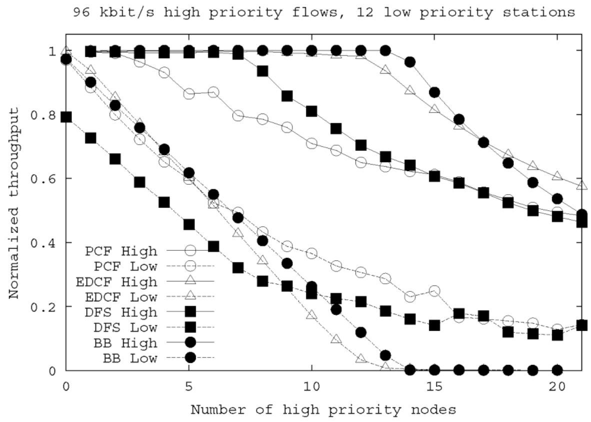
- Scheduler at AP
  - decide when to poll the stations (CF-Poll)
  - decide what parameters should be assigned to each TC (AIFS, CW, P)
- Queue State Indicator
  - measurements of arrival rates at client
  - notify the scheduler at AP
- Traffic Specification
  - based on RSVP or other higher-layer protocol
  - MAC-layer negociation

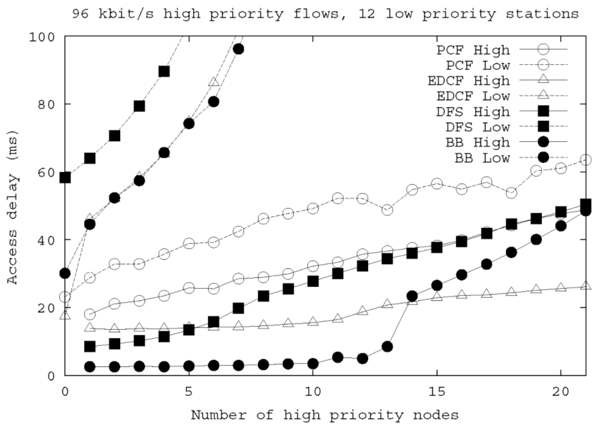
Other proposals and evaluation

Blackburst

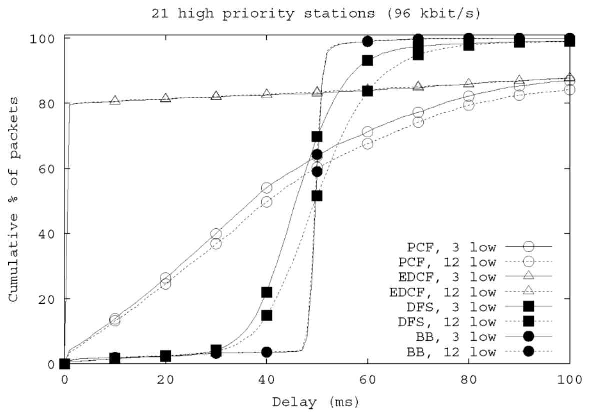


- An access method for high priority traffic
  - low priority traffic: DCF mode with CSMA/CA
  - high priority voice traffic: wait PIFS and compete for channel
  - voice users transmit "black bursts" jam of length proportional to waiting time, wait OIFS (Observation IFS)
  - voice user who waited longest wins (longest black burst)
  - voice connections tend to evenly spread out in time frame
- Sobrinho, Krishnakumar, Globecom 96

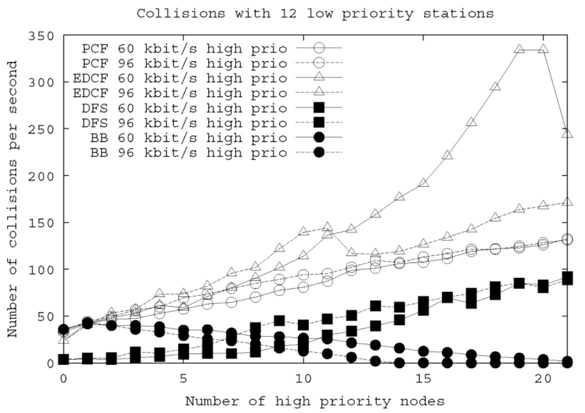




79



80



81

Evaluation (Lindgren et al.)

- Blackburst
  - gives the best performance to high priority traffic both with regard to throughput and delay
  - requires constant access intervals
  - at high loads, low priority traffic is starved
- EDCF
  - good performance to high priority traffic, low average delay
  - at high loads, low priority traffic is starved, long tail of the delay distribution
  - suffers from high collision rate
- DFS
  - relative differentiation
  - fair share - low priority traffic is not starved

82

Conclusions

- 802.11e has several advantages
  - EDCF can provide differentiated access among different user priorities
  - EDCF Bursting increases throughput performance at the cost of larger delays for voice and reduces contention overhead
  - backward compatibility: 802.11 stations operate as best effort
- Will it be of any use?
  - 802.1D priorities are not widely used on wired LANs
  - need a sophisticated scheduler at AP
  - how to prevent a station from using the highest priority?
- Provide a good wireless MAC instead
  - optimize for high throughput, low latency (good short term fairness), low collision rate
  - equal contention windows for fairness
  - should be adaptive - optimal contention windows depend on the load (number of contending stations)

83

References and further reading

- D. Kitchin, "The 802.11 MAC Protocol & Quality of Service"
- S. Mangold et al. "IEEE 802.11e Wireless LAN for Quality of Service"
- H. Zhu et al., "A Survey of Quality of Service in IEEE 802.11 Networks", IEEE Wireless Comm.
- A. Lindgren et al. "Evaluation of Quality of Service Schemes for IEEE 802.11Wireless LANs"

84