



Trasmissione adattativa di dati multimediali su IP

Davide Quaglia
a.a. 2006/2007

1

Sommario

- Def. del problema e cenni storici
- Transmission Control Protocol (TCP)
- Caratteristiche del traffico multimediale
 - coesistenza con TCP
- Obiettivo dei protocolli adattativi
- Architettura di un sistema adattativo
- Problematiche
 - tipo di rete
 - reattivita'
 - impatto sulla qualita' percepita
- Esempi di realizzazioni

2

Cenni storici e definizione del problema

3

Cenni storici

- Le reti telefoniche sono tradizionalmente a commutazione di circuito
 - Il canale e' di capacita' nota e costante lungo tutto il circuito
 - controllo di ammissione
 - Il canale e' riservato all'utente anche in caso di inutilizzo
 - Non serve alcuna forma di adattamento
- Internet
 - interconnessione di reti eterogenee
 - no controllo di ammissione
 - no garanzie di Qualita' del Servizio
 - traffico dati: caratteristica trasmissione a burst

4

Definizione del problema

- Capacita' effettiva del canale
 - capacita' intrinseca (velocita' dell'interfaccia)
 - presenza di traffico interferente
 - presenza di malfunzionamenti (wireless)
 - varia nel tempo
 - limitata dal tratto peggiore (dipende dal routing)
- Wireless: improvvisi malfunzionamenti di tratti di rete
 - interferenza
 - path loss
 - fading e shadowing

5

Definizione del problema (2)

- Trasmettere ad un rate maggiore della capacita' effettiva del canale comporta
 - aumento dei ritardi di transito
 - perdita di pacchetti
- Congestioni nelle code dei router

6

Una prima soluzione: Transmission Control Protocol (TCP)

7

TCP

- Creazione di connessioni su IP (che e' non connesso)
 - byte-oriented (come un file)
 - affidabili (senza perdite)
 - consegna ordinata dei dati
- Ritrasmissione dei dati non confermati
- Adatta il rate di trasmissione
 - alla capacita' del ricevitore (controllo di flusso)
 - alla capacita' della rete (congestion avoidance)
- Inizia a trasmettere lentamente e poi aumenta il rate se non avvengono perdite

8

ACK

- Ogni pacchetto TCP e' chiamato segmento
- Sequence number: ultimo byte che si sta trasmettendo
- Ack number: numero del prossimo byte atteso
- RTO: retransmission time-out
- Il meccanismo di ACK ha un doppio significato
 - conferma di ricezione
 - unito al meccanismo di *stop-and-go* permette di controllare il numero di byte che stanno transitando sulla rete
 - nuovi byte possono entrare quando vecchi byte escono

9

Stop-and-go

- Il trasmettitore puo' trasmettere solo un certo numero di byte prima di ricevere il corrispondente ACK
- Due parametri:
 - Finestra di ricezione (controllo di flusso)
 - Finestra di congestione
- Il valore di tali parametri cambia durante la vita della connessione
- Viene usato il minore dei due parametri

10

Controllo di flusso

- Finestra di ricezione
 - indica il numero di byte che il TX puo' mandare senza aspettare l'ACK
 - indicato dal RX per non venire sovraccaricato
 - indicato nel messaggio di ACK

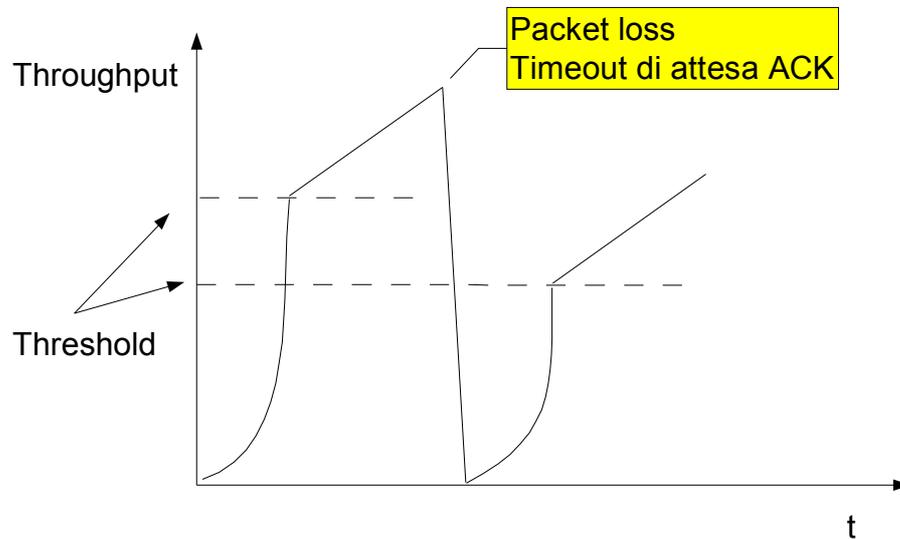
11

Controllo di congestione

- Congestion window *cwin*
 - indica il numero di byte che il trasmettitore puo' mandare senza aspettare l'ACK
 - mantenuta dal trasmettitore
- Slow start phase: inizialmente $cwin = S$ (1 segmento)
- *cwin* viene incrementata di S per ogni segmento di cui si e' ottenuta la conferma
 - crescita esponenziale fino ad un valore detto *threshold*
 - dopo il valore di *threshold* si ha crescita lineare
- Dopo un time-out di ACK *cwin* viene rimesso a S e *threshold* viene dimezzato

12

Profilo di throughput



13

Throughput dopo il transitorio

$$T = \frac{S}{R\sqrt{\frac{2p}{3}} + t_{RTO}\sqrt{\frac{3p}{8}}p(1+32p^2)}$$

- S: dimensione in byte del segmento
- R: round trip time
- t_{RTO} : time-out di ritrasmissione
- p: packet loss rate

14

Conseguenze

- Le sorgenti TCP si dividono equamente la banda
- Ritardi dovuti all'attesa di ACK
- Throughput fluttuante
- Basso utilizzo del canale in presenza di molte perdite
- Possibile non equita' in reti con elevato round trip time

15

Caratteristiche del traffico multimediale

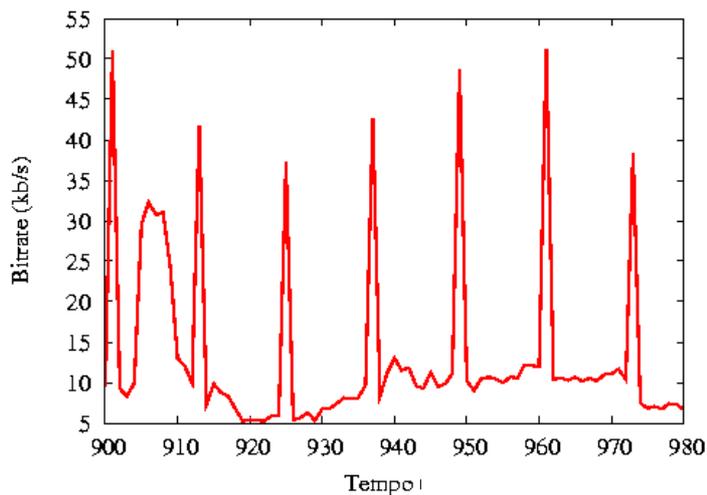
16

Caratteristiche

- Trasmissione real-time
- Tollerante alle perdite ma non ai ritardi e alle variazioni dei ritardi
- Bitrate di sorgente predefinito dallo schema di compressione
 - costante
 - variabile ma periodico (MPEG)
- Requisiti opposti al normale traffico dati che utilizza TCP !!!
- Si preferisce utilizzare UDP

17

Traffico VBR di MPEG



18

Coesistenza TCP e UDP

- Flussi TCP tendono a dividersi equamente la capacita' del canale
- Un flusso TCP riduce il rate di trasmissione quando perde pacchetti (time-out di ACK)
- Un flusso UDP “base” non riduce il rate di trasmissione quando si perdono pacchetti
- I flussi UDP aggrediscono i flussi TCP che
 - finiscono per usare solo la capacita' residua
 - si spengono (starvation)

19

Trasmissione adattativa di flussi multimediali

20

Obiettivi

- utilizzo efficiente delle risorse
 - stima della capacita' del canale
- reattivita' alle variazioni di capacita' del canale
 - congestione
 - failure
- evitare le congestioni -> comportamento equo
 - verso TCP (TCP-friendliness)
 - traffico dotato di uguale algoritmo (fairness)
- robustezza e stabilita'
 - variazioni di rate della sorgente
 - variazioni di capacita' del canale
 - perdita di messaggi di controllo

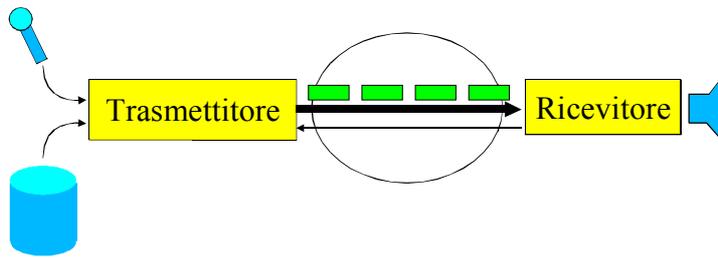
21

Architettura di un sistema adattativo

- Sistema ad anello chiuso
- Stima delle condizioni del canale
 - raccolta statistiche
 - elaborazione statistiche
- Trasmissione dei messaggi di feedback
- Algoritmo di controllo
 - relazione tra condizioni del canale e rate di trasmissione
- Sistema di variazione del rate di trasmissione

22

Ipotesi semplificativa



- Si considera come unità base la coppia TX/RX
- Nei sistemi interattivi ci sono 2 coppie
- I messaggi che chiudono l'anello vanno a gravare sul traffico nel senso opposto

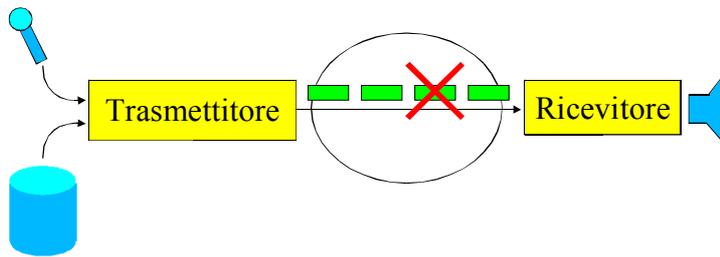
23

Qualità percepita al ricevitore

- bitrate di compressione
- ritardo subito dai pacchetti
- variazione del ritardo subito dai pacchetti
- percentuale di pacchetti persi
- burstiness delle perdite

24

Perdita di pacchetti



- perdita alle code dei router
- errori sui bit non recuperabili
- arrivo in ritardo al ricevitore

25

Perdita di pacchetti (2)

- Rilevazione
 - numero progressivo nell'header del pacchetto multimediale
 - checksum o CRC per rilevare (non corregg.) errori sui bit
- Misura
 - Packet loss rate (PLR) su una certa finestra temporale

$$PLR = \frac{\text{numero pacchetti persi}}{\text{numero pacchetti spediti}}$$

- Finestra temporale piccola --> valore istantaneo
- Finestra temporale grande --> valore medio

26

Stima della perdita di pacchetti

- il ricevitore invia un ACK/NACK su singolo pacchetto o frame (pesante)
- il ricevitore invia un valore di packet loss rate su una finestra temporale (RTCP)
- su richiesta del trasmettitore

27

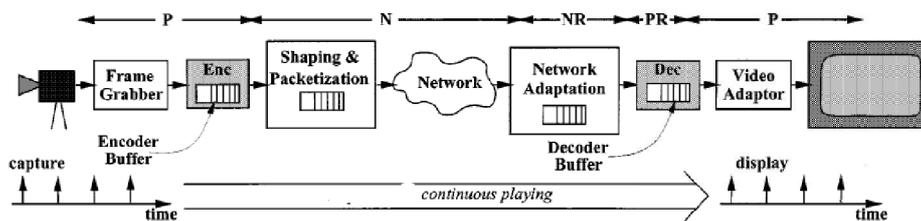
Ritardo end-to-end

- Def: tempo intercorrente tra l'istante in cui la grandezza fisica (luce, suono) viene campionata e l'istante in cui il corrispondente campione viene riprodotto al ricevitore
- Al fine di una riproduzione fluida il ritardo end-to-end dovrebbe essere costante.

28

Ritardo end-to-end (2)

- Componenti
 - compressione (ritardo algoritmico + tempo di elaboraz.)
 - accodamento prima della trasmissione
 - propagazione
 - accodamento ed elaborazione nei router
 - accodamento prima della decompressione
 - decompressione (tempo di elaborazione)



29

Stima del ritardo end-to-end

- Al ricevitore non è possibile conoscere il ritardo end-to-end in maniera assoluta a causa della mancanza di un clock comune tra TX e RX
- Stima dalla distanza temporale tra i pacchetti
 - R_i istante di arrivo del pacchetto i -esimo
 - T intervallo tra la riproduzione di 2 pacchetti consecutivi

$$D_i = R_i - R_{i-1} - T$$

- idealmente $D_i = 0$
- in pratica: valor medio \sim nullo

30

Valor medio di D_i

- Calcolo esatto
 - Sia $\{D_i\}$ l'insieme dei ritardi misurati su N pacchetti

$$\mu_D = \frac{\sum_1^N D_i}{N} \approx 0$$

- Stima effettuata “al volo” al ricevitore

$$\bar{D}_i = \alpha D_i + (1 - \alpha) \bar{D}_{i-1} \quad 0 \leq \alpha \leq 1 \quad \alpha = \frac{1}{8}$$

31

Variatione di D_i

- Calcolo esatto (deviazione standard)
 - Sia $\{D_i\}$ l'insieme dei ritardi misurati su N pacchetti

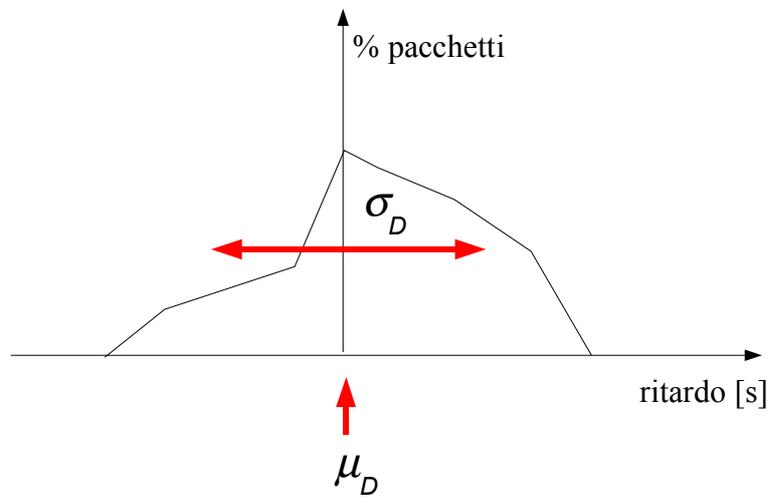
$$\sigma_D = \sqrt{\frac{\sum_1^N (D_i - \mu_D)^2}{N}} \approx \sqrt{\frac{\sum_1^N D_i^2}{N}}$$

- Stima effettuata “al volo” dal ricevitore

$$J_i = \alpha |D_i| + (1 - \alpha) J_{i-1} \quad 0 \leq \alpha \leq 1 \quad \alpha = \frac{1}{16}$$

32

Distribuzione dei D_i



33

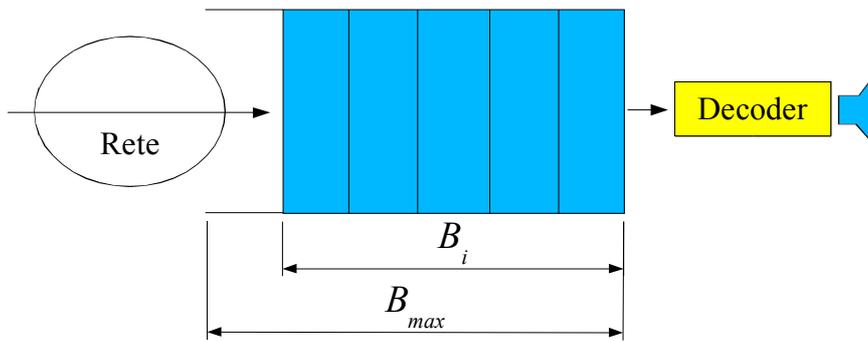
Jitter

- **Def: variazione del ritardo end-to-end**
- **Cause: variazione nel tempo delle condizioni del canale di comunicazione**
 - variazioni del traffico totale nella rete
 - variazioni della qualità del segnale radio
 - movimento del trasmettitore e/o ricevitore
 - presenza di ostacoli
 - attenuazione del segnale con la distanza
 - movimento di altri nodi che vanno ad interferire con trasmettitore e/o ricevitore
- **Effetti**
 - ritardo elevato (vedi slide precedenti) e timeout
 - arrivo “a valanga” dei pacchetti

34

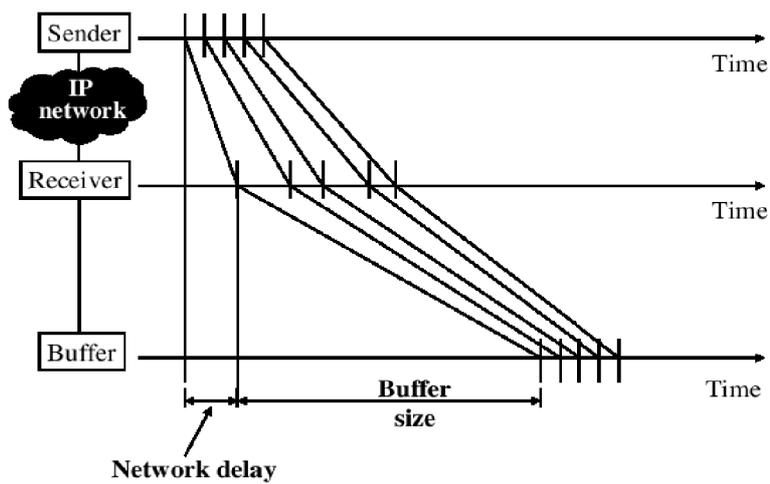
Buffer anti-jitter

- Coda di pacchetti al ricevitore
- Trasforma le variazioni di ritardo in un ritardo fisso
 - Aumenta il ritardo end-to-end (compromesso)



35

Buffer anti-jitter (2)



36

Buffer anti-jitter (3)

- All'istante di decodifica del pacchetto i -esimo
 - Se vale la relazione

$$D_i = R_i - R_{i-1} - T \geq B_i$$

- allora il pacchetto è considerato perso per timeout e svuotamento del buffer (buffer underflow)
- altrimenti se

$$B_i - D_i > B_{max}$$

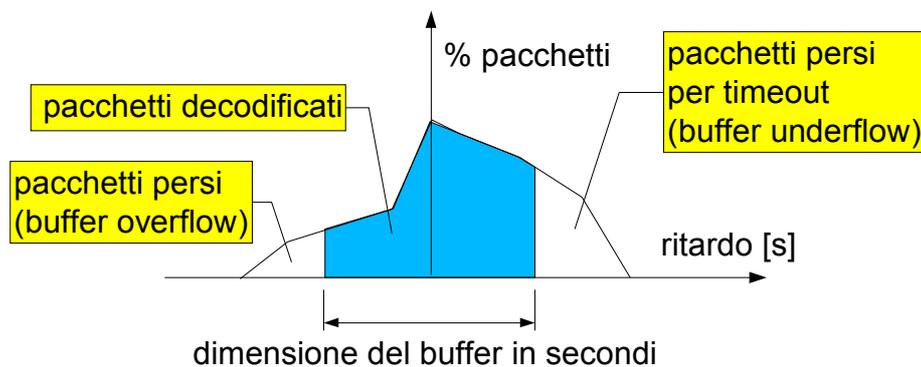
- allora il pacchetto è considerato perso per riempimento del buffer (buffer overflow)
- altrimenti il pacchetto è decodificato

37

Buffer anti-jitter (4)

- L'occupazione del buffer all'istante i -esimo

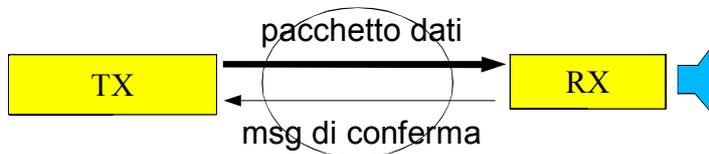
$$B_{i+1} = \min(B_{max}, \max(0, B_i - D_i))$$



38

Round trip time

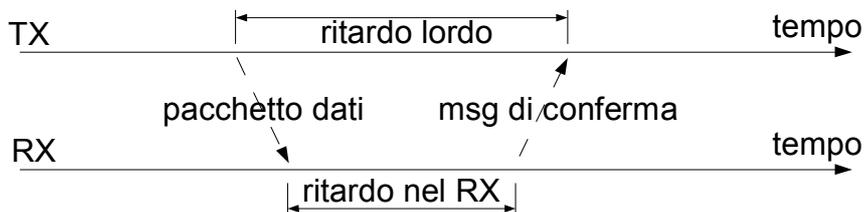
- Tempo intercorrente tra l'istante di trasmissione di un pacchetto dati e la ricezione di un messaggio di conferma (acknowledge).
- Viene calcolato al TX
- Il messaggio di conferma può essere messo in un
 - pacchetto specifico
 - pacchetto informativo sulle statistiche di ricezione
 - pacchetto multimediale trasmesso nella direzione opposta (applicazioni interattive)



39

Calcolo del round trip time

- Viene fatto al TX
- Il messaggio di conferma dal RX deve riportare il ritardo nel RX

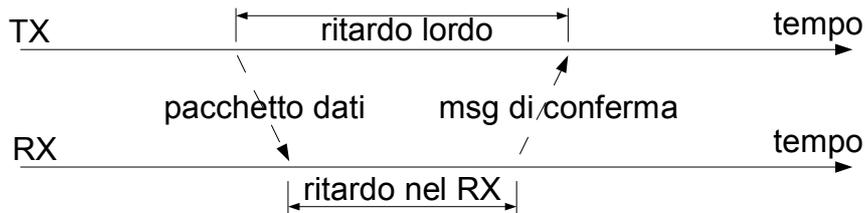


$$RTT = \text{ritardo lordo} - \text{ritardo nel RX}$$

40

Stima del ritardo tramite round trip time

- Metodo di stima del ritardo di **propagazione nella rete** mediante il calcolo del round trip time
 - Viene fatto al TX (a differenza del metodo del D_p)
 - Il messaggio di conferma dal RX deve riportare il ritardo nel RX



$$\delta = \frac{\text{ritardo lordo} - \text{ritardo nel RX}}{2}$$

41

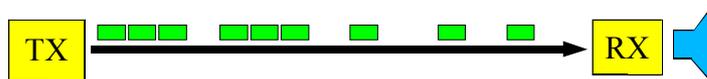
Stima del ritardo tramite round trip time (2)

- Possibile errore di stima se i pacchetti fanno percorsi diversi nelle due direzioni oppure la capacità nelle due direzioni è diversa
- Es:
 - satellite per il download + linea terrestre per l'upload
 - ADSL
 - GPRS (3*9600 b/s download, 1*9600 b/s upload)

42

Burstiness di arrivo

- lunghezza media (su un intervallo di osservazione) delle sequenze di pacchetti aventi inter-packet gap nullo o sotto una certa soglia

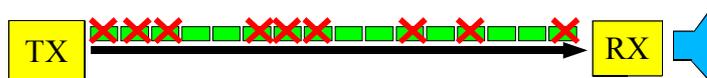


$$\text{lunghezza media del burst} = \frac{(3+3+1+1+1)}{5}$$
$$= \frac{9}{5} = 1.8$$

43

Burstiness delle perdite

- lunghezza media (su un intervallo di osservazione) delle sequenze di pacchetti persi adiacenti



$$\text{lunghezza media del burst} = \frac{(3+3+1+1+1)}{5}$$
$$= \frac{9}{5} = 1.8$$

44

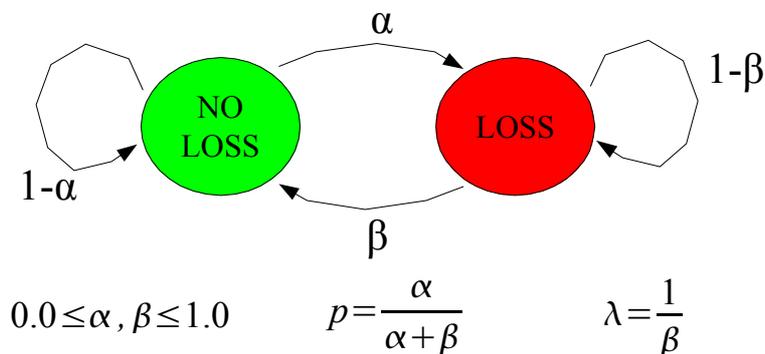
Elaborazione delle statistiche

- Può avvenire al trasmettitore che riceve le statistiche dal ricevitore (RTCP)
- Può avvenire al ricevitore che comanda direttamente al trasmettitore la variazione del rate di trasmissione
- Può avvenire al ricevitore che estrapola un modello astratto del canale e trasmette i parametri al trasmettitore
 - modello di Gilbert-Elliot

45

Canale con perdita con memoria

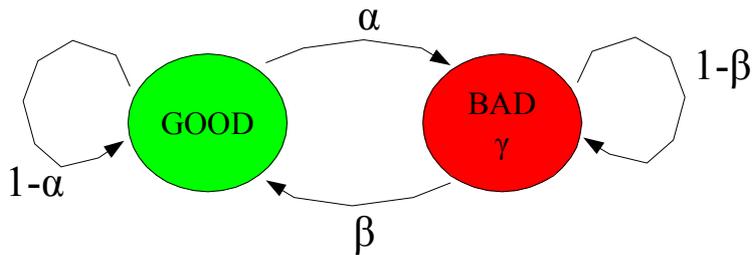
- eventi di perdita statisticamente dipendenti
- presenza di burst di pacchetti persi di lunghezza media λ
- Catena di Markov a due stati



46

Modello di Gilbert-Elliot

- Catena di Markov gerarchica
- Nello stato GOOD: nessun errore (o perdita)
- Nello stato BAD: errori (o perdite) indipendenti con probabilità γ



47

Messaggi di feedback

- ACK/NACK su singolo pacchetto o frame (pesante)
- packet loss rate e inter-packet gap mediati su una finestra temporale
 - RTCP Receiver Report
- comandi di modifica del rate di trasmissione

48

Algoritmo di controllo

- relazione tra condizioni del canale e rate di trasmissione
- algoritmi Additive Increase Multiplicative Decrease (AIMD)
 - derivato dal TCP
 - buona TCP-friendliness e fairness
- algoritmi basati su formule per il calcolo della capacità effettiva del canale

49

Sistemi AIMD

- Multimedia over TCP (HTTP)
 - facile da implementare
 - risolve problemi di NAT e firewall
 - Es: Icecast
- Rate Adaptation Protocol (RAP)
 - Layered coding
- Selective Retransmission RTP (SR-RTP)
- Stream Control Transmission Protocol (SCTP) - RFC 2960
- Datagram Congestion Control Protocol (DCCP) – vedere: draft-ietf-dccp-spec-13

50

Sistemi AIMD: pro e contro

- Convivono bene con traffico simile e TCP
- Come il TCP non si prestano bene alla natura del multimedia

51

Sistemi basati su formule

- TCP Friendly Rate Control (TFRC)
- Video Transport Protocol (VTP)
- Datagram Congestion Control Protocol (DCCP) – vedere: draft-ietf-dccp-spec-13

52

TCP Friendly Rate Control

- Additive increase
- Decrease according to the estimated capacity
- Allowed capacity estimated in terms of packet size, loss event rate, RTT, and the retransmission timer value
- No congestion window
- No slow start phase
- Equation is based on steady state and doesn't model transient behavior
- Slow reaction to sudden increase of bandwidth

$$T = \frac{S}{R \sqrt{\frac{2p}{3}} + t_{RTO} \sqrt{\frac{3p}{8}} p (1 + 32 p^2)}$$

53

Sistema di adattamento della sorgente

- quantizzazione variabile
- coding mode selection (variazione tra I,P,B)
- codifica a livelli (layered, Fine Grain Scalability - FGS)
- aggiunta di FEC
- lunghezza dei pacchetti
- lunghezza del buffer di playout
- rate di trasmissione variabile (Variable Time Scale Streaming - VTSS)

54

Codificatori vocali a bitrate variabile

- Speex (opensource Ogg Project)
 - bitrate disponibili (kb/s):
 - 2.15 - 24.6 (narrowband),
 - 4.0 - 44.2 (wideband)
- ETSI GSM-AMR (Adaptive multirate)
 - bitrate disponibili (kb/s):
 - 4.75 – 12.2 (narrowband)
 - 6.6 – 23.85 (wideband)
- ISO MPEG-4 CELP

55

Compromessi nell'adattativita'

- ridondanza
- ritardo algoritmico (dipende dall'algoritmo di compressione, dalla lunghezza dei pacchetti e dalla lunghezza del buffer di playout)
- occupazione di banda
- consumo di energia (dispositivi mobili)
- qualita' dovuta al rate di compressione

56

Impedimenti alla reattivita'

- Valore di round trip time
- perdita di pacchetti di feedback/comandi
- calcolo di medie necessarie per evitare fluttuazioni brusche
- periodo di feedback (da barattare con occupazione di banda)
- possibilita' di adattamento della sorgente solo in certi istanti (GOP)

57

Adattativita' per il multicast

- Adattamento della sorgente sulla base dei feedback dei vari ricevitori
 - Adattamento al feedback peggiore (penalizzante)
 - Adattamento ad un feedback medio (penalizzante per alcuni e rischioso per altri)
- Receiver-driven Layered Multicast (McCanne, 1997)
 - Layered source coding
 - Each layer is sent over a different multicast group
 - Each receiver adjusts its reception rate by joining and leaving multicast groups

58

Problemi

- Canali asimmetrici: ADSL, GPRS
 - la presenza di capacita' diverse nelle due direzioni puo' far diminuire l'utilizzo del canale ?
- Wireless
 - la percentuale di pacchetti persi non dovuta a congestioni ma a problemi radio e' considerevole
 - si possono creare falsi allarmi di congestione che fanno abbassare l'utilizzo del canale