

5. Gestione e controllo del traffico ATM

Classificazione delle Categorie di Servizio ATM

Relazione fra le Classi di Servizio e le Categorie di Servizio

Parametri di definizione del traffico e QoS

Peak Cell Rate e CDVT

Caratteristiche di una connessione ATM (I.371; TM4.0)

Funzioni e procedure per l'accettazione, il controllo del traffico e della congestione: CAC, UPC/NPC, GCRA

Esempio di applicazione di GCRA con traffico costante e a burst

Policing con valutazione delle priorità

Definizione della conformità della categorie di servizio secondo l'ATM Forum

Controllo del flusso ABR

Struttura delle celle RM

1



La tecnologia ATM supporta, come già accennato, servizi multimediali di diversa natura con garanzia di qualità. Per poter mantenere fede a questa "promessa" è fondamentale il ruolo svolto dai meccanismi di gestione del traffico e della congestione. In ambito di standardizzazione internazionale l'ATM Forum per primo ha redatto un documento che disciplina tale attività. Tale documento si intitola Traffic Management Specification v4.0. L'ITU-T ha fatto altrettanto con la raccomandazione I.371 (Traffic Control and Congestion Control in B-ISDN).

Lo scopo di questi documenti è quello di specificare le procedure ed i parametri per poter applicare il controllo del traffico e della congestione.

In primo luogo il controllo del traffico serve a proteggere la rete ed i sistemi terminali da eventuali congestioni che porterebbero a non avere i livelli di qualità voluti.

Successivamente mediante il controllo della congestione si cerca di eliminare le cause che determinano il degrado della qualità dovute all'utilizzo improprio delle risorse della rete.

L'ATM Forum e l'ITU-T hanno pertanto definito un insieme di categorie di servizio che presentano peculiari combinazioni di parametri di traffico e di qualità.

In particolare, l'ATM Forum ha definito le seguenti categorie di servizi:

1. Servizi in tempo reale

Tasso costante (Constant Bit Rate, CBR)

Tasso variabile con vincoli di tempo reale (real-time Variable Bit Rate, rt-VBR)

2. Servizi non in tempo reale

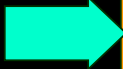




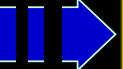
Tasso variabile e assenza di vincoli di tempo reale (non-real-time Variable Bit Rate, nrt-VBR)

Tasso disponibile (Available Bit Rate, ABR)

Tasso non specificato (Unspecified Bit Rate, UBR)

L'ATM fornisce quindi servizi sia in tempo reale che non in tempo reale. Una rete basata su ATM può infatti sopportare un'ampia gamma di tipi di traffico, fra cui si può citare il traffico derivato da flussi sincroni TDM che può appoggiarsi sul servizio a tasso costante (Constant Bit Rate, CBR), il traffico in tempo reale generato dalla compressione di voce e video, che può far uso del servizio a tasso variabile in tempo reale (real-time Variable Bit Rate, rt-VBR), il traffico con specifici requisiti di qualità, ma non in tempo reale, che può far uso del servizio VBR non in tempo reale (non-real-time Variable Bit Rate, nrt-VBR), ed infine il traffico IP, che può usare il servizio a tasso disponibile (Available Bit Rate, ABR) e quello a tasso non specificato (Unspecified Bit Rate, UBR).

Classificazione delle Categorie di Servizio ATM

ATM FORUM		ITU-T	
Traffic Management Specification 4.0		Traffic Control and Congestion Control in B-ISDN I.371	
ATM Service Category		ATM Transfer Capability	
	Constant Bit Rate (CBR)		Deterministic Bit Rate (DBR)
	Real Time Variable Bit Rate (rt-VBR)		<i>(for further study)</i>
	Non Real Time Variable Bit Rate (nrt-VBR)		Statistical Bit Rate (SBR)
	Available Bit Rate (ABR)		Available Bit Rate (ABR)
	Unspecified Bit Rate (UBR)		<i>(no equivalent)</i>
	<i>(no equivalent)</i>		ATM Block Transfer (ABT)



Le molteplici tipologie di connessioni che possono presentarsi in una rete ATM non possono essere discriminate solo in base al valore di certi parametri di traffico o di QoS, ma possono richiedere l'impiego di parametri diversi e l'applicazione di funzioni di controllo diverse. Le diverse classi di trasferimento ATM sono state quindi introdotte per raggruppare le connessioni a cui può essere applicata la stessa descrizione parametrica e lo stesso insieme di funzioni di controllo.

Sono state finora definite dall'ITU-T (nella I.371) e dall'ATM Forum (nella TM 4.0) cinque diverse classi:

- una classe dedicata a connessioni a banda costante, o comunque soggette ad allocazione di risorse basata sulla frequenza di picco, detta CBR (constant Bit Rate) nella terminologia ATM Forum e DBR (Deterministic Bit Rate) in quella ITU-T;
- una classe relativa a connessioni a banda variabile, con caratteristiche prevedibili ma non modificabili in funzione dello stato di occupazione della rete, e le cui esigenze di QoS richiedono una allocazione preventiva delle risorse, anche su base statistica; questa classe è detta VBR (variable Bit Rate) dall'ATM Forum e SBR (Statistical Bit Rate) dall'ITU-T.
- una classe dedicata a connessioni a banda variabile, con caratteristiche difficilmente prevedibili ma prive di particolari requisiti di QoS, detta UBR (Unspecified Bit Rate); questa classe è, per ora, prevista solo dall'ATM Forum;
- una classe dedicata a connessioni a banda variabile le cui caratteristiche possono essere modificate in funzione dell'occupazione della rete, indicata come ABR (Available Bit Rate);
- una classe dedicata a connessioni che fanno uso di procedure di negoziazione delle risorse su base burst oltre che su base chiamata, detta ABT (ATM Block Transfer); questa classe è oggi prevista solo dall'ITU-T.

Per quanto riguarda la QoS negoziata tra la TM 4.0 e la I.356 ci sono alcune differenze:

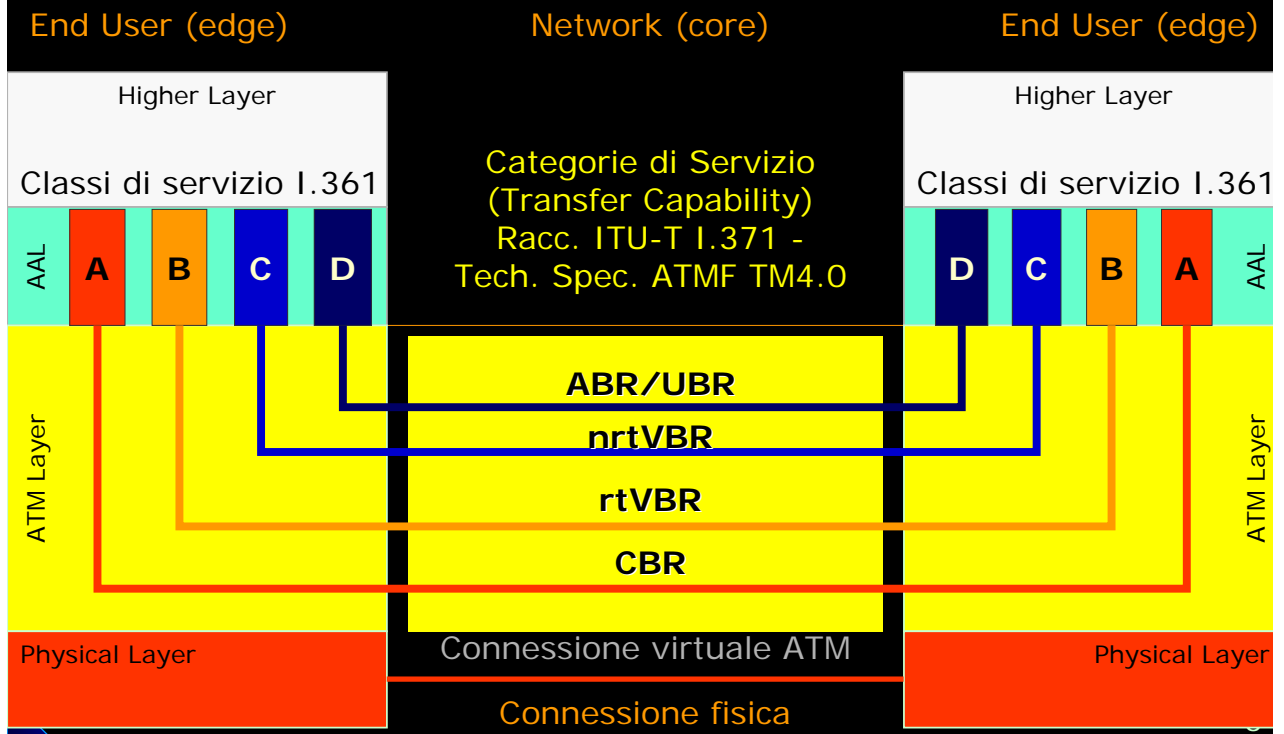
L'ITU-T prevede la suddivisione in classi: Classe 1 (stringent); Classe 2 (tolerant); Classe 3 (bi-level); Classe U (unspecified).

L'ATM Forum prevede invece la negoziazione di singoli parametri di qualità: Peak-to-peak Cell Delay Variation; Maximum Cell Transfer Delay; Cell Loss Ratio.

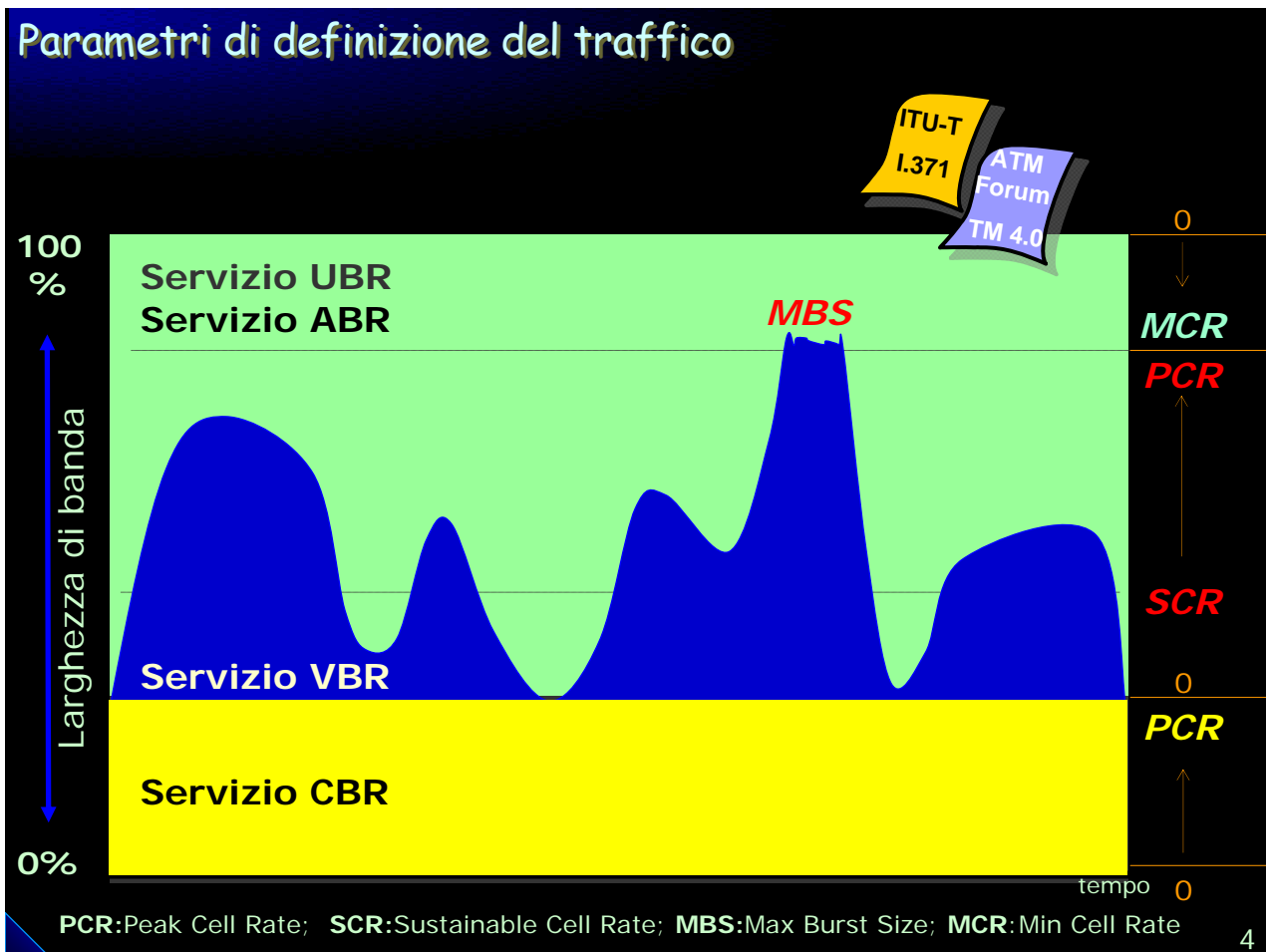
Degli aspetti di qualità se ne parlerà approfonditamente nell'apposita sezione.

Nella presente documentazione si fa riferimento alla terminologia dell'ATM Forum in quanto, essendo la prima ad essere stata introdotta, è la più utilizzata.

Relazione fra le Classi di Servizio e le Categorie di Servizio



Note:



I parametri che descrivono il traffico ATM, gli STD (Source Traffic Descriptor) che sono stati definiti in ITU-T I.371 sono la frequenza di picco di cella, caratterizzata dal parametro PCR (Peak Cell Rate), la frequenza media di cella, caratterizzata invece dai due parametri SCR (Sustainable Cell Rate) e MBS (Maximum Burst Size) che, congiuntamente considerati, oltre a definire la frequenza media, limitano la lunghezza massima del burst che una sorgente può trasmettere alla frequenza di picco (PCR).

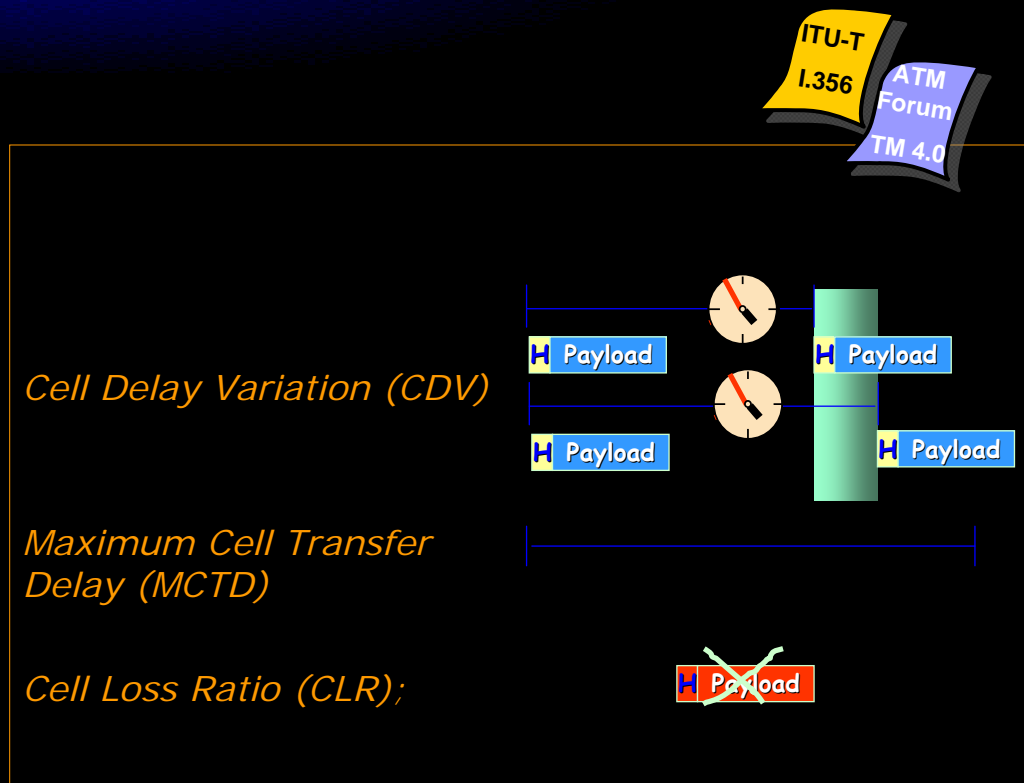
Il PCR limita la frequenza massima con cui una connessione può trasmettere celle: esso è definito come il reciproco del tempo minimo di inter-arrivo (cioè della differenza dei corrispondenti istanti d'arrivo) tra due richieste consecutive di trasmissione di celle di una stessa connessione.

SCR e MBS limitano la frequenza media con cui una connessione può trasmettere celle, specificando il valore nominale di essa attraverso il parametro SCR ed indicando attraverso il parametro MBS, l'anticipo massimo con cui una cella può essere trasmessa rispetto al caso ideale di trasmissione con banda costante, pari a SCR.

Il parametro MCR rappresenta il valore minimo garantito.

La figura rappresenta il modo in cui una rete alloca le risorse in una situazione stazionaria (nessuna apertura o chiusura di canali virtuali).

Parametri della Qualità del Servizio



5



Mediante la caratterizzazione della QoS (Qualità del Servizio) la rete oltre a modulare l'allocazione delle risorse trasmissive sulle necessità effettive delle connessioni, dà delle priorità a quelle con caratteristiche più stringenti.

La differenza più importante tra diverse applicazioni riguarda il valore e la variabilità del ritardo, denominata jitter, tollerati. Le applicazioni in tempo reale, generalmente, riguardano il trasferimento di un flusso di informazioni verso un utente, finalizzato ad una sua riproduzione a destinazione che deve apparire come se il flusso fosse riprodotto nella sorgente. Per esempio, un utente si aspetta che un flusso di informazioni video o audio giunga in modo continuo, e la mancanza di continuità o un tasso di perdita eccessivo si riflette in un rilevante degrado della qualità. Anche le applicazioni che comportano interazioni tra persone fisiche sono soggette a stretti vincoli sul ritardo, ritardo che diventa percettibile quando assume valori di qualche centinaia di millisecondi. In tali circostanze i requisiti richiesti alla rete ATM, per la commutazione e la consegna dei dati in tempo reale, sono molto impegnativi.

Per le applicazioni interattive è quindi importante che l'MCTD sia contenuto e che il CDV sia costante.

Per le applicazioni orientate al trasferimento dei dati bisogna che la rete offra un'alta affidabilità garantendo un CLR molto piccolo. Si vedrà in seguito la caratterizzazione delle categorie di servizio in base ai parametri di traffico e di qualità secondo l'ATM Forum.

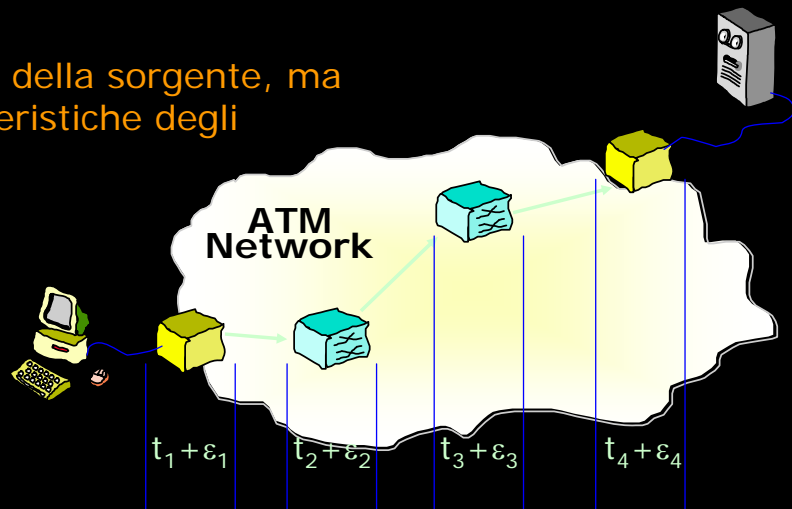
In generale, nel caso di servizi non in tempo reale, categoria cui appartengono quelle applicazioni che hanno caratteristiche di traffico intermittente e che non hanno vincoli sul ritardo, la rete può usufruire di una maggiore flessibilità nel trattare il traffico e può utilizzare al meglio il multiplexing statistico per aumentare la propria efficienza.

Tolleranza del ritardo di trasferimento (CDVT)

Esprime il valore massimo della variazione del ritardo di trasferimento delle celle (CDV) introdotto da:

- apparati di utente;
- apparati della rete.

Non è un parametro della sorgente, ma dipende dalle caratteristiche degli apparati



6

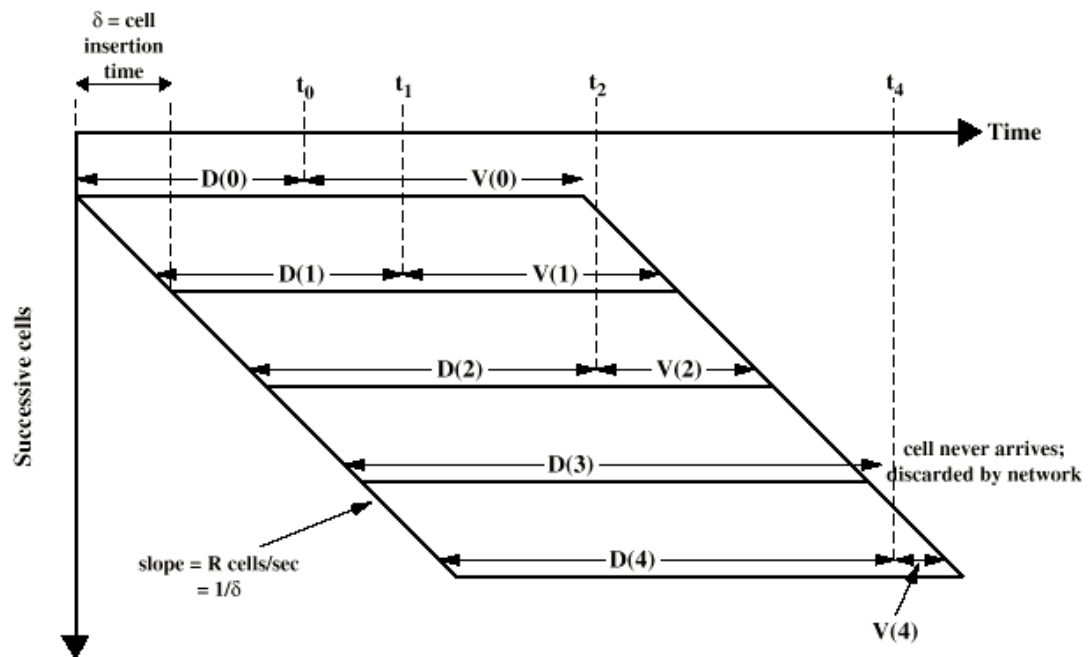


La caratterizzazione di una connessione ATM deve essere completata specificando la tolleranza alla variazione del ritardo (CDVT) relativa alla frequenza di picco che quella relativa alla frequenza media.

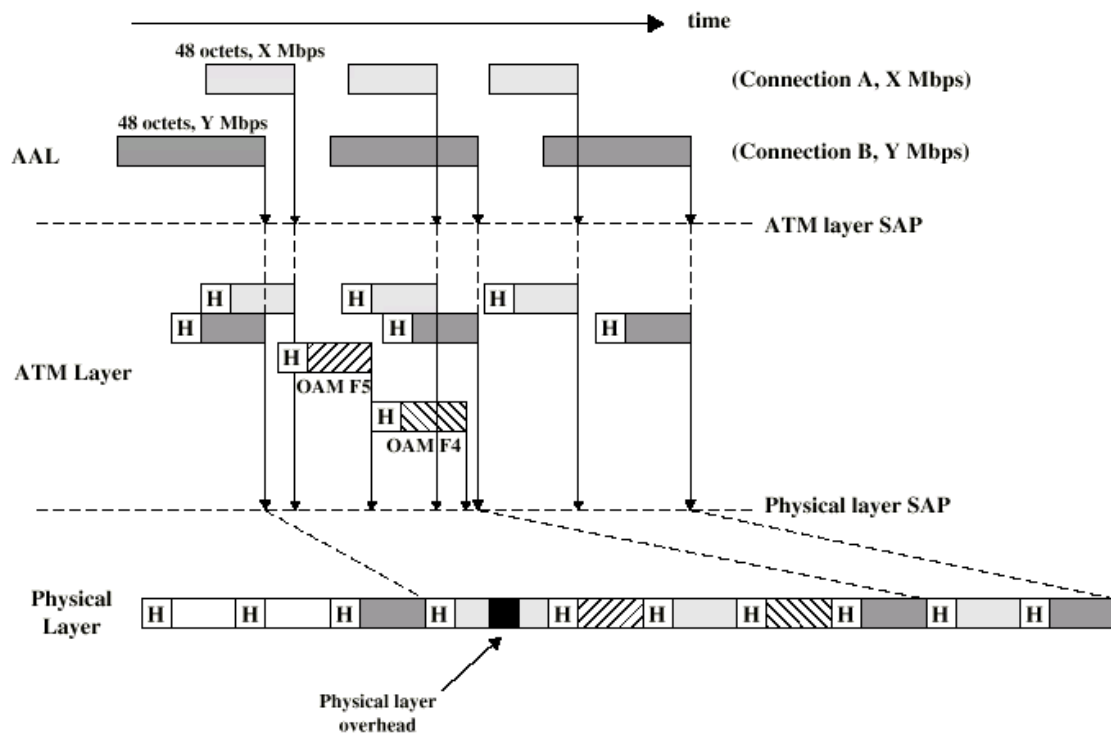
Per il momento non è ancora stato chiarito se la CDVT debba essere la stessa nei due casi o se possa invece assumere valori diversi; la scelta oggi prevalente prevede l'utilizzo di due valori diversi che possono essere dichiarati esplicitamente o implicitamente nel contratto di traffico.

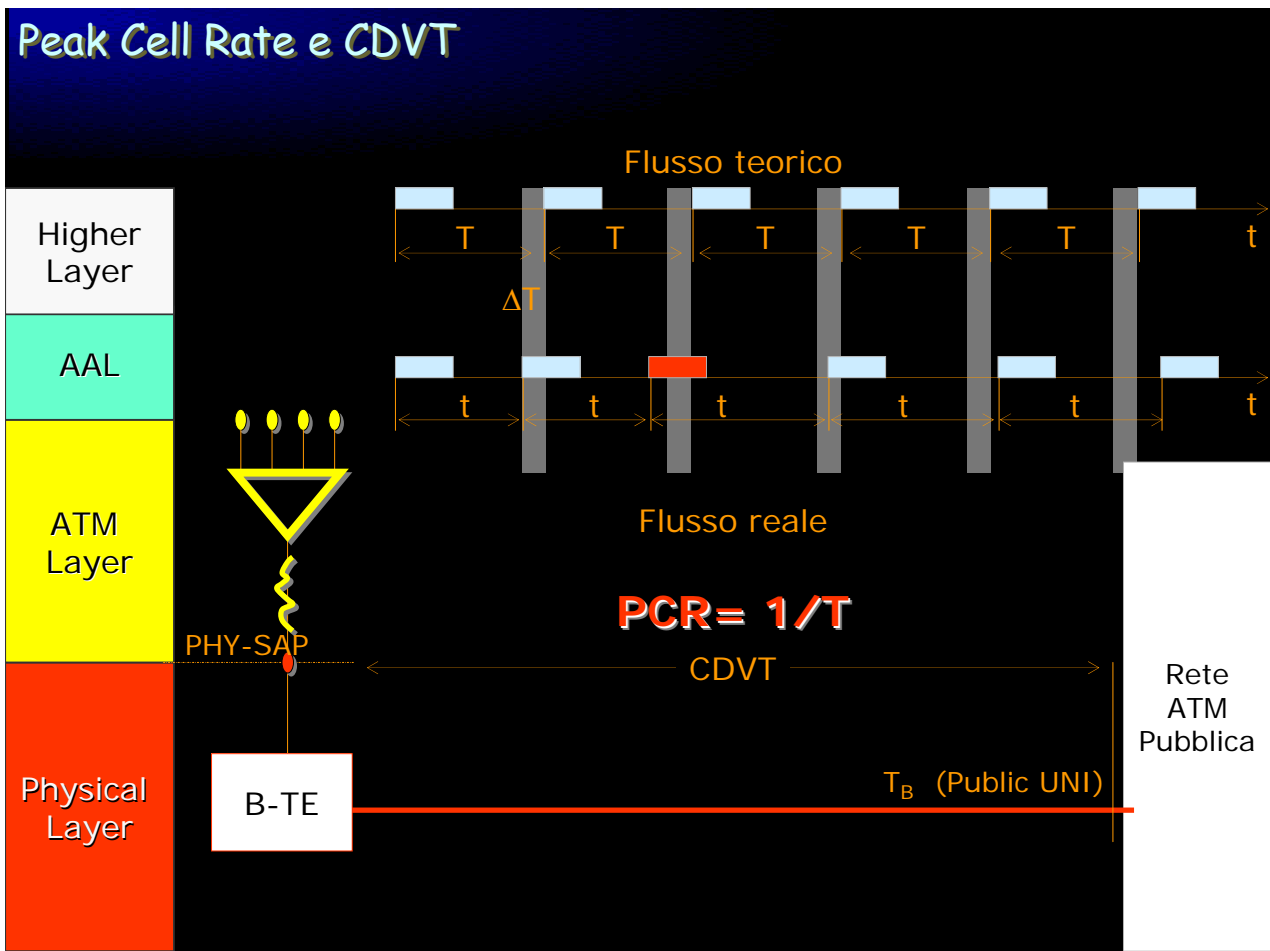
Le variazioni introdotte dagli apparati d'utente o dalla rete su questi parametri sono espresse mediante la CDVT (Cell Delay Variation Tolerance), che rappresenta in pratica il limite superiore di queste variazioni. La CDVT non è, quindi, un parametro della sorgente ma dipende dalle caratteristiche degli apparati d'utente e delle reti attraversate dalla connessione, e cambia lungo il percorso della connessione stessa.

Time Re-assembly of CBR Cells



Origins of Cell Delay Variation





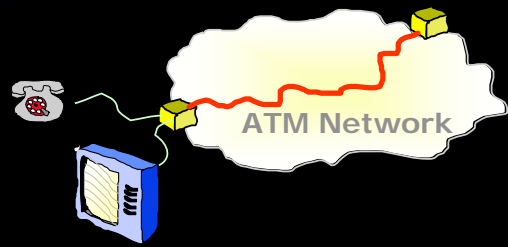
Il parametro di traffico PCR è uno di quelli importanti che caratterizzano la sorgente.

Il PCR di una connessione virtuale indica la massima velocità con cui si può trasferire il traffico, e di conseguenza, richiede alla rete la risorsa massima di banda da assegnare alla connessione.

Con riferimento al modello di terminale in figura, la PCR viene valutata al PHY-SAP, cioè nel punto in cui le celle vengono formate, a valle della moltiplicazione e dello shaping, per essere trasmesse a livello fisico.

Al valore di PCR deve sempre essere associato un valore di tolleranza CDVT per permettere alla rete di allocare opportunamente le proprie risorse, come ad esempio la dimensione dei buffer. In base a questa coppia di valori è inoltre possibile, durante il trasferimento sulla connessione attivata, effettuare il controllo del traffico (policing).

Constant Bit Rate



Peak Cell Rate

Sustainable Cell Rate

Maximum Burst Size

Minimum Cell Rate

CDVT

Cell Loss Ratio

Cell Delay Variation

Cell Transfer Delay

10

Servizi:

- video e fonia tradizionali
- emulazione di circuito

Caratteristiche del trasferimento

- basso ritardo
- bassa variazione del ritardo

Traffic
ParameterQoS
Parameter

Il servizio CBR è forse quello più semplice da definire e viene utilizzato da applicazioni che richiedono un tasso di trasmissione dati fisso e continuamente disponibile per tutta la durata della connessione, ed un limite massimo relativamente piccolo sul ritardo di trasmissione. Il CBR è comunemente usato per informazioni audio/video non compresse; alcuni esempi a riguardo sono: Video conferenza, Audio interattivo (ad esempio, telefonia), Distribuzione audio/video (ad esempio, televisione, apprendimento a distanza, video a pagamento), Recupero audio/video (ad esempio, video su richiesta (video on demand), biblioteca audio).

Questa classe mette a disposizione delle singole connessioni una quantità di banda statica pari alla frequenza massima di celle PCR negoziata; è stata definita principalmente per soddisfare le esigenze di servizi isocroni ed a banda costante, ma non è limitata a questi impieghi. Prevede la caratterizzazione del traffico mediante il solo parametro PCR. Le connessioni sono soggette ad un controllo di accettazione che prevede l'allocazione della banda di picco. Il parametro PCR è soggetto anche a una verifica di conformità da parte della rete. Le funzioni di controllo del traffico rilevanti per questa classe sono perciò la funzione di accettazione delle connessioni e la funzione di controllo di conformità relative al parametro PCR.

La classe CBR permette di supportare servizi ad emulazione di circuito. A questa categoria appartengono sia la fonia che la trasmissione video; a tali servizi viene garantito un ritardo ed una variazione del ritardo piuttosto contenuti.

Parametri del traffico: di interesse solo il PCR (Peak Cell Rate).

Qualità del servizio: sensibile a CDV (Cell Delay Variation) e MCTD (Maximum Cell Transfer Delay).

Applicazioni: videoconferenza, audio interattivo, audio/video-distribuzione (TV, Pay Per View,...), Video on Demand, ecc...

Real-Time Variable Bit Rate

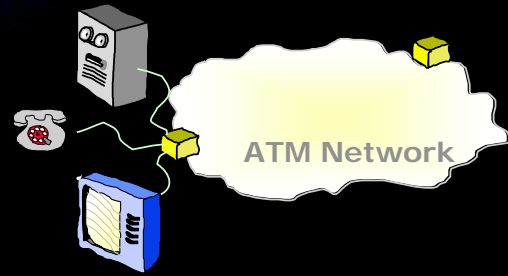


Servizi:

- comunicazioni interattive di nuova generazione

Caratteristiche del trasferimento

- ritardo contenuto
- possibilità di perdita di celle



Peak Cell Rate

Sustainable Cell Rate

Maximum Burst Size

Minimum Cell Rate

CDVT

Traffic
Parameter

QoS
Parameter

Cell Loss Ratio

Cell Delay Variation

Cell Transfer Delay

11



La categoria Real-Time Variable Bit Rate (rt-VBR) fa parte di quelle applicazioni che possiedono dei requisiti in termini di tempo, ovvero, quelle applicazioni che richiedono un ritardo e una variazione di ritardo molto contenute. La differenza principale tra le applicazioni adatte ad un servizio rt-VBR e quelle adatte a uno CBR consiste nel fatto che le applicazioni rt-VBR trasmettono con un tasso variabile nel tempo; in altre parole, una sorgente rt-VBR è di tipo impulsivo. Per esempio, l'approccio standard alla compressione video si traduce in una sequenza di quadri di immagini che hanno dimensioni diverse. Poiché il video in tempo reale richiede un tasso di trasmissione di quadri d'immagini uniforme, il tasso di emissione dei dati varia nel tempo.

Il servizio rt-VBR permette alla rete una maggiore flessibilità rispetto al CBR, poiché la rete può moltiplicare un certo numero di connessioni usando la stessa capacità richiesta da un singolo servizio CBR, pur offrendo il servizio richiesto ad ogni connessione.

La categoria rt-VBR rappresenta i futuri servizi interattivi che sfruttano la natura a burst di tutte le comunicazioni. In questo caso diventa importante il contenimento del ritardo, ma non ci sono particolari problemi sulla perdita di celle.

Parametri del traffico: PCR (Peak Cell Rate), SCR (Sustainable Cell Rate) e MBS (Maximum Burst Size).

Qualità del servizio: sensibile a CDV (Cell Delay Variation) e MCTD (Maximum Cell Transfer Delay).

Applicazioni: VBR in generale è adatto per ogni applicazione supportata dalla moltiplicazione statistica lato terminale. In particolare il real-time è adatto per voce e video a pacchetto ATM che sfruttano nuovi algoritmi di compressione.

Non-Real-Time Variable Bit Rate



Servizi:

- comunicazioni non interattive (es. e-mail)

Caratteristiche del trasferimento

- ritardi notevoli
- nessuna perdita di celle



Traffic
Parameter

Peak Cell Rate

Sustainable Cell Rate

Maximum Burst Size

Minimum Cell Rate

CDVT

QoS
Parameter

Cell Loss Ratio

Cell Delay Variation

Cell Transfer Delay

12



Nel caso di alcune applicazioni non in tempo reale, è possibile caratterizzare il flusso di traffico previsto, in modo tale che la rete possa offrire una qualità di servizio migliore (QoS) anche nelle zone in cui si subiscono ritardi o perdite; questo genere di applicazioni possono utilizzare il servizio Non-Real-Time Variable Bit Rate (nrt-VBR). Con questo servizio il sistema terminale specifica un tasso di picco e un tasso medio (di cella) sostenibili e una misura dell'intermittenza delle celle o del modo in cui le celle sono raggruppate. Con queste informazioni la rete può fare in modo di assicurare delle risorse per offrire un ritardo relativamente basso e una perdita minima di celle.

Il servizio nrt-VBR può essere utilizzato per il trasferimento di dati che hanno delle criticità in termini di tempi di risposta. Alcuni esempi sono le prenotazioni aeree, le transazioni bancarie e la supervisione di processo.

nrt-VBR è complementare alla categoria rt-VBR. In questo caso, non interessa particolarmente il ritardo, ma si richiede che nessuna cella venga persa. La posta elettronica (e-mail) è un esempio di questo tipo di traffico.

Parametri del traffico: PCR (Peak Cell Rate), SCR (Sustainable Cell Rate) e MBS (Maximum Burst Size).

Qualità del servizio: sensibile a MCTD (Maximum Cell Transfer Delay) e CLR (Cell Loss Ratio).

Applicazioni: VBR in generale è adatto per ogni applicazione supportata dalla multiplexazione statistica lato terminale. In particolare il non real-time è adatto per data transfer, frame relay interworking, applicazioni che sfruttano processi di transazione non interattivi (servizi bancari, servizi di prenotazione, consultazioni banche dati, ecc.).

Unspecified Bit Rate



Servizi:

- comunicazioni dati a basso costo
- servizi connection-less

Caratteristiche del trasferimento

- ritardi notevoli
- mancanza di garanzia sull'arrivo dei dati

Traffic
Parameter

Peak Cell Rate

Sustainable Cell Rate

Maximum Burst Size

Minimum Cell Rate

CDVT

QoS
Parameter

Cell Loss Ratio

Cell Delay Variation

Cell Transfer Delay

13



In un istante qualsiasi, una certa quantità della capacità complessiva di una rete ATM è impegnata nel trasportare il traffico CBR e i due tipi di VBR. La capacità residua rimane disponibile per una o entrambe le seguenti ragioni: 1) non tutte le risorse complessivamente disponibili sono state impegnate dal traffico CBR e VBR; 2) la natura del traffico VBR porta talvolta all'utilizzo di una capacità inferiore rispetto a quella impegnata e tutta quella che rimane libera potrebbe essere resa disponibile per il servizio UBR. Questo servizio è adatto a quelle applicazioni che possono tollerare ritardi variabili e perdite di celle, come in genere accade per il traffico basato su TCP. Con l'UBR, le celle sono inviate con politica first-in-first-out (FIFO, il primo arrivato è il primo ad uscire), utilizzando la capacità non impegnata dagli altri servizi; pertanto sono possibili sia perdite di celle che ritardi. Nei confronti di una sorgente UBR non viene effettuata alcuna negoziazione e non viene fornito alcun riscontro su eventuali congestioni, ossia si ha un servizio normalmente chiamato best-effort (al meglio). Alcuni esempi di applicazioni UBR sono: Trasferimento di dati, immagini e testo, messaggistica, distribuzione, recupero d'informazioni; Terminale remoto.

Questa classe non prevede alcuna forma di allocazione delle risorse alle diverse connessioni, ma si basa su una condivisione libera delle risorse disponibili. Essa è stata introdotta per offrire un servizio economico per il trasporto con gli attuali protocolli di trasmissione dati (tipicamente il TCP), che offrono un traffico difficilmente prevedibile (in quanto dipende dalla perdita di pacchetti che si verifica nella rete), ma sono relativamente robusti rispetto alla perdita di celle, non hanno stringenti requisiti temporali e posseggono una certa capacità di adattare la velocità di trasmissione allo stato presunto della rete. In questo caso è prevista una caratterizzazione del traffico mediante il solo parametro PCR e non è offerta alcuna garanzia di QoS. Le connessioni sono soggette ad un controllo di accettazione piuttosto blando, finalizzato essenzialmente a limitare il numero di connessioni di questo tipo che possono essere presenti contemporaneamente, in modo da consentire un'accettabile condivisione delle risorse. Il parametro PCR potrebbe essere soggetto a verifica di conformità da parte della rete. È importante sottolineare che questa classe è per ora considerata solo dall'ATM Forum, e non dall'ITU-T. In quest'ultimo ambito è infatti prevalsa per il momento la tendenza a considerare il modo UBR non tanto come una classe di trasferimento, quanto come un servizio da offrire associando la classe DBR (equivalente della CBR) ad una qualità del Servizio di tipo U.

Available Bit Rate



Servizi:

- comunicazioni dati a basso costo
- servizi connection-less

Caratteristiche del trasferimento

- ritardi notevoli

Traffic
Parameter

Peak Cell Rate

Sustainable Cell Rate

Maximum Burst Size

Minimum Cell Rate

CDVT

QoS
Parameter

Cell Loss Ratio

Cell Delay Variation

Cell Transfer Delay

14



Applicazioni impulsive, che utilizzano un protocollo affidabile end-to-end, come il TCP, possono determinare la presenza di congestione nella rete osservando un aumento del ritardo complessivo sul percorso di andata e di ritorno e la perdita di alcuni pacchetti, come vedremo in seguito. Tuttavia, il TCP non ha un meccanismo che permette di capire se le risorse sulla rete sono divise equamente tra più connessioni dello stesso tipo, e, inoltre, il TCP non minimizza la congestione nel modo più efficiente possibile, come potrebbe fare se utilizzasse delle informazioni esplicite provenienti dai nodi congestionati della rete. Per migliorare il servizio fornito alle sorgenti impulsive, che altrimenti utilizzerebbero UBR, è stato definito il servizio ABR. Un'applicazione che utilizza ABR specifica il tasso di picco di emissione delle celle (Peak Cell Rate, PCR) che vorrebbe utilizzare, e quello minimo (Minimum Cell Rate, MCR) che richiede. La rete assegna le risorse in modo tale che tutte le applicazioni ABR ricevano almeno la capacità MCR richiesta, mentre quella non utilizzata viene condivisa in maniera equa e controllata da tutte le sorgenti ABR. Il meccanismo ABR usa un riscontro esplicito verso le sorgenti per assicurare che la capacità venga assegnata in modo equo, mentre le risorse non utilizzate dalle sorgenti ABR rimangono disponibili per il traffico UBR. Un esempio di un'applicazione che dovrebbe utilizzare l'ABR è l'interconnessione di LAN, in cui i sistemi terminali collegati ad una rete ATM sono i router.

La Classe di Trasferimento ABR è stata pensata come miglioramento della classe UBR, e dovrebbe essere utilizzabile dai protocolli di trasmissione dati della prossima generazione, ovvero da quelli appositamente definiti per il trasporto in tecnica ATM. In linea di principio la classe ABR consente di garantire una bassa probabilità di perdita delle celle, a patto che le sorgenti modulino la loro frequenza di trasmissione di celle in funzione dei segnali di controllo generati dalla rete. La classe ABR prevede che il traffico venga caratterizzato specificando il valore massimo di banda richiesta, mediante il parametro PCR e, eventualmente, anche il valore minimo MCR.

Questa classe prevede la capacità di variare dinamicamente le risorse allocate a ciascuna connessione, adattandole alle reali esigenze dei terminali e, soprattutto, allo stato di occupazione della rete. Per realizzare questa modalità di funzionamento è necessaria l'utilizzazione di un anello di controllo basato su un protocollo di RM (Resource Management), le cui unità dati sono trasportate a livello ATM all'interno di una connessione controllata.

Parametri del traffico: PCR (Peak Cell Rate), MCR (Minimum Cell Rate).

Qualità del servizio: sensibile a CLR (Cell Loss Ratio).

Applicazioni: Interconnessione di LAN, LAN Emulation, servizi dati a costi contenuti, ecc.

Le funzioni di controllo del traffico rilevanti per questa classe sono ancora in fase di definizione.

Caratteristiche di una connessione ATM (I.371; TM 4.0)

Occorrono delle regole precise per la caratterizzazione delle connessioni e la gestione del traffico:



La caratterizzazione delle singole connessioni permette alla rete di conoscere le informazioni necessarie ad allocare le risorse ed a stabilire quali procedure di controllo debbano essere adottate per garantire la Qualità del Servizio richiesta.

La caratterizzazione di una connessione virtuale (Virtual Path o Virtual Channel) comprende perciò l'indicazione della Categoria di trasferimento (ATC), cioè del tipo di servizio di livello ATM, e la specifica nell'ambito di questa classe dei valori dei parametri di traffico e della QoS richiesta.

Durante la fase di accettazione della connessione, viene negoziato il traffico specificando i seguenti parametri:

La classe di trasferimento di livello ATM (ATC: ATM Transfer Capability);

Il descrittore del traffico generato dalla sorgente STD (Source Traffic Descriptor) che comprende tutti i parametri atti a caratterizzare il traffico generato dai diversi flussi indipendenti compresi nella connessione; questi flussi, e di conseguenza la struttura dello STD, dipendono dall'ATC usata per la connessione.

Uno dei parametri fondamentali è il **CDVT** (Cell Delay Variation Tolerance) relativo ai diversi parametri descritti nell'STD, da applicare all'interfaccia sulla quale è definito il contratto di traffico.

La Qualità del Servizio associata alla connessione;

L'indicazione se considerare o meno applicabile l'operazione di tagging, cioè la forzatura ad 1 del valore del bit Cell Loss Priority delle celle ad alta priorità non conformi al contratto di traffico. Questa indicazione è significativa solo per il modo di trasferimento Statistical Bit Rate (SBR), adatto al trasporto di applicazioni che generano un traffico a frequenza di cifra variabile;

La definizione di conformità del traffico con cui verrà valutato durante il servizio.

L'insieme dei parametri di traffico **STD** e la **definizione di conformità** dichiarate per ciascun flusso definisce il **descrittore di traffico della connessione ATM**.



Le reti ATM devono allocare le risorse della rete in modo flessibile ed ottimale. In particolare, la rete deve supportare vari tipi di traffico e fornire livelli di servizio diversi.

Per esempio, come precedentemente dimostrato, la voce richiede un basso ritardo ed una bassa variazione del ritardo. La rete deve quindi allocare le proprie risorse per garantire questa condizione. Il concetto usato per risolvere questo tipo di problemi viene chiamato gestione del traffico.

Quando una connessione (channel o path) viene attivata, viene anche specificato un opportuno contratto del traffico. Questo permette alla rete o agli operatori della rete ATM di esaminare l'utilizzo esistente e determinare se un collegamento può essere stabilito con le condizioni richieste. Se le risorse della rete non sono disponibili, la connessione può essere negata. La rete permetterà la connessione solo se trova un percorso dotato di risorse sufficienti per garantire le caratteristiche del traffico richieste.

Funzioni e procedure per l'accettazione, il controllo del traffico e della congestione

Connection Admission Control (CAC)

Usage/Network Parameter Control (UPC/NPC)

Network Resource Management (NRM)

Priority control

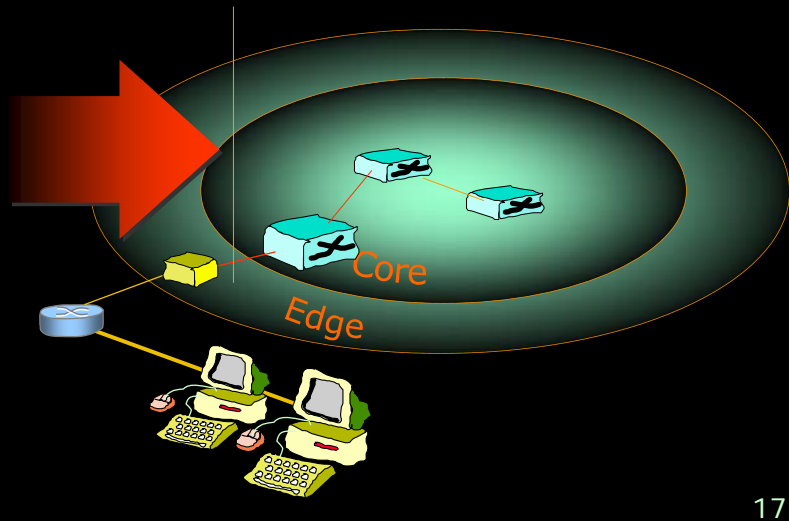
Traffic Shaping

Frame Discard

ABR Flow Control

EFCI Control

Feedback control



17



Le funzioni che la rete svolge per garantire la qualità del servizio sono le seguenti:

Connection Admission Control (CAC) - è un set di azioni preso dalla rete nella fase di set up per verificare se la connessione richiesta può essere accettata (in tal caso va ricercato il VC che soddisfa la QoS richiesta) oppure deve essere respinta;

Usage/Network Parameter Control (UPC/NPC) - è un set di azioni preso dalla rete per monitorare e verificare la validità del traffico, in base ai parametri precedentemente negoziati per la connessione, nel punto di accesso alla rete. Il suo scopo è quello di proteggere la rete da comportamenti "illeciti" che possono pregiudicare la QoS delle connessioni già attivate. La funzione UPC può quindi prendere provvedimenti per fare rientrare il traffico entro i parametri concordati mediante lo scarto o la diminuzione della priorità delle celle. L'UPC viene svolta all'interfaccia UNI (User Access), mentre NPC all'interfaccia tra due reti (Network Access). Tale funzione viene anche definita Traffic Policing.

Priority control - per alcuni servizi la sorgente può generare dei flussi di celle a diversa priorità (bassa o alta) in funzione dell'assegnazione del bit CLP. Se la rete dà significato a questo bit può selettivamente scartare le celle dando precedenza a quelli di bassa priorità.

Traffic Shaping - lo shaping può essere usato per modificare la forma del traffico;

Network Resource Management (NRM) - prevede il provisioning delle risorse disponibili, viene costantemente eseguito dalla rete;

Frame Discard - per eliminare una congestione si opera lo scarto di celle, ma per avere un'azione efficace esiste un meccanismo che permette lo scarto anche di tutte le altre celle appartenenti alla stessa struttura PDU di livello superiore;

ABR Flow Control - permette un colloquio tra i vari nodi di una connessione per utilizzare in modo adattativo le risorse di banda. E' un meccanismo di gestione delle risorse dell'ATM Layer: mediante l'uso di celle RM si possono modificare le caratteristiche della connessione richiesta;

Explicit Congestion Forward Indication - permette di informare il destinatario della cella che una condizione di congestione è avvenuta in un nodo attraversato. Questo permette a protocolli di livello superiore di prendere dei provvedimenti mediante una riduzione della Cell Rate.

Feedback control - è un set di azioni preso dalla rete e dai sistemi terminali per regolare il traffico sulla connessione attivata;

Connection Admission Control

Informazioni valutate da CAC

Categoria di servizio

Source Traffic Descriptor

Tolleranze rispetto ai valori richiesti

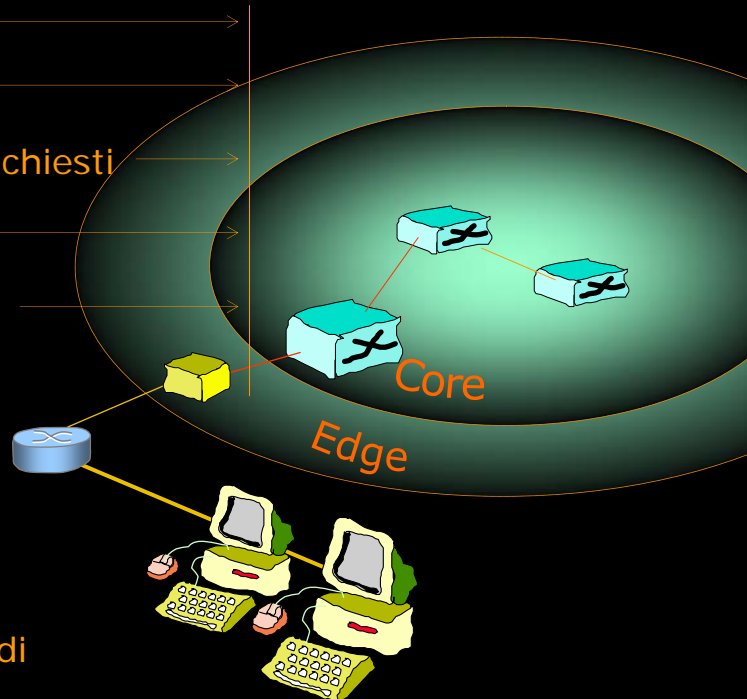
Qualità del servizio

Definizione di una conformità

Azioni con riscontro positivo

Assegnazione del percorso e delle risorse di rete

Assegnazione dei parametri di conformità UPC/NPC



18



Il contratto del traffico rappresenta quella funzione precedente all'invio delle informazioni che regola il flusso dei dati accettati dalla rete.

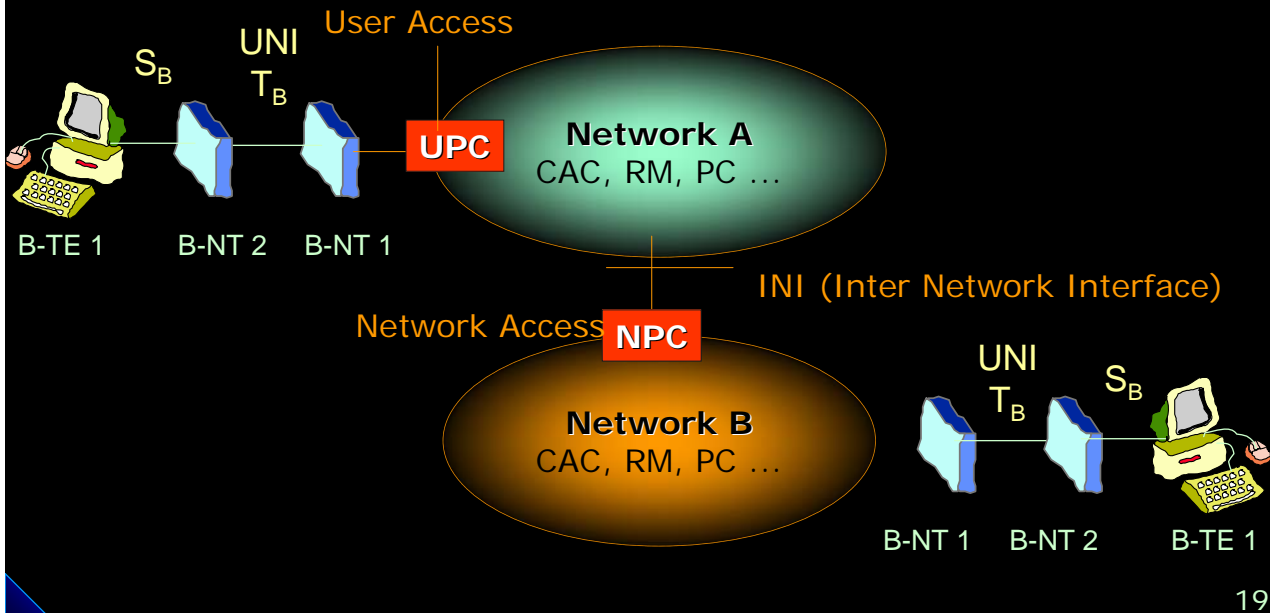
La prima cosa che viene fatta dalla rete è l'accettazione in base alle risorse disponibili. La rete pertanto si impegna a fornire le risorse in modo affidabile, in base al QoS e al profilo di traffico concordate, l'utente deve però generare del traffico conforme al contratto.

Tipicamente il contratto del traffico avviene all'interfaccia UNI, cioè tra l'utente finale ed un nodo d'accesso della rete pubblica. All'interno delle reti private non vengono date delle precise indicazioni sul profilo di traffico consentito.

Anche nei punti di interconnessione tra diversi gestori viene concordato un contratto di traffico. A differenza dell'interfaccia UNI non si entra nel dettaglio delle singole connessioni virtuali, ma del volume di traffico totale scambiato.

Dopo l'accettazione la rete svolge delle funzioni di routing per determinare il percorso migliore in base alle risorse, inoltre assegna i parametri di conformità ai nodi che svolgono le funzioni di controllo UPC/NPC.

Usage/Network Parameter Control



19



Una volta che la funzione di controllo di ammissione delle connessioni accetta una connessione, la funzione di controllo dei parametri di utilizzazione (Usage Parameter Control, UPC) controlla se il traffico generato è conforme a quello dichiarato nel contratto di traffico. Lo scopo principale del controllo dei parametri di utilizzazione è quello di proteggere le risorse di rete da un sovraccarico generato da una connessione, che influenzerebbe negativamente la QoS di altre connessioni, rilevando le violazioni dei parametri assegnati e prendendo contromisure adeguate. Il controllo dei parametri di utilizzazione può essere fatto sia a livello di canale virtuale, sia a livello di percorso virtuale. Fra i due, il più importante è il controllo a livello di VPC, perché le risorse di rete sono solitamente assegnate su base percorso virtuale, con la capacità condivisa tra i canali virtuali che costituiscono il percorso. Il controllo dei parametri di utilizzazione ha due scopi principali:

- Controllare il tasso di picco delle celle (PCR) e la variazione del ritardo di cella (CDV);
- Controllare il tasso di cella sostenibile (SCR) e la tolleranza agli impulsi ad esso associata (MBS).

La funzione svolta dall'UPC/NPC viene anche detta **Policing** del traffico (sorveglianza). In base ai risultati dell'algoritmo di conformità GCRA (Generic Cell Rate Algorithm) di tipo Leaky Bucket applicato a tutte le celle di ogni connessione, vengono presi dei provvedimenti che possono essere:

cella conforme --> trasmissione

cella non conforme --> scarto

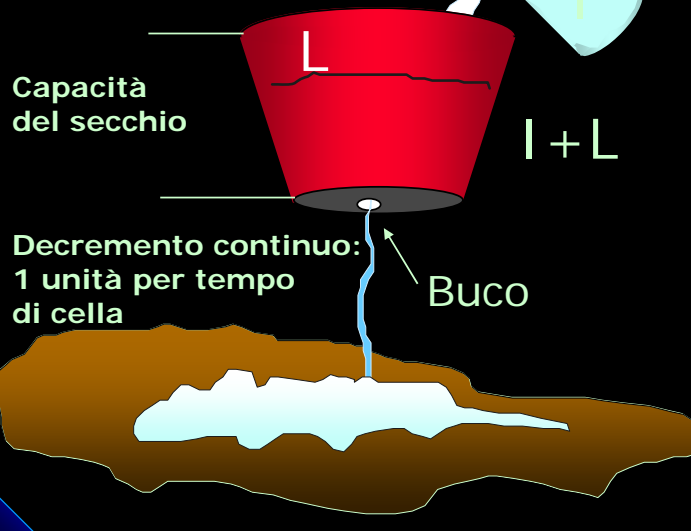
cella non conforme --> tagging, diminuzione della priorità

Generic Cell Rate Algorithm (GCRA)



Leaky Bucket

Liquido versato quando viene congegiata una cella appartenente alla connessione



— Controlla se il Cell Rate in una data connessione è conforme a quanto stipulato dal contratto del traffico.

— Usa l'algoritmo GCRA che prevede due parametri:

— Incremento I ;

— Limite L .

— La funzione del GCRA è analoga a quella svolta da un secchio bucato nel quale si versano frequentemente delle quantità di liquido.

— Il contratto del traffico viene violato quando il secchio trabocca.

20



Una delle tecniche per il controllo del traffico è quella basata sul "Generic Cell Rate Algorithm" (GCRA); tale algoritmo specifica in dettaglio quali sono le condizioni in cui un flusso di celle viola il contratto del traffico; esso è definito in base a due parametri: il parametro dell'incremento " I " e il parametro del limite " L ".

Il GCRA può essere associato per analogia ad un secchio bucato che perde del liquido in modo continuo (Leaky Bucket). Un'unità di liquido fuoriesce dal secchio per ogni intervallo di cella (inverso del Peak Cell Rate), che essendo di lunghezza fissa equivale ad un tempo ben determinato in relazione all'interfaccia trasmissiva utilizzata.

Il parametro " I " tiene sotto controllo il Cell Rate del flusso in esame, mentre " L " controlla la dimensione dei burst (si vedrà più avanti con un esempio).

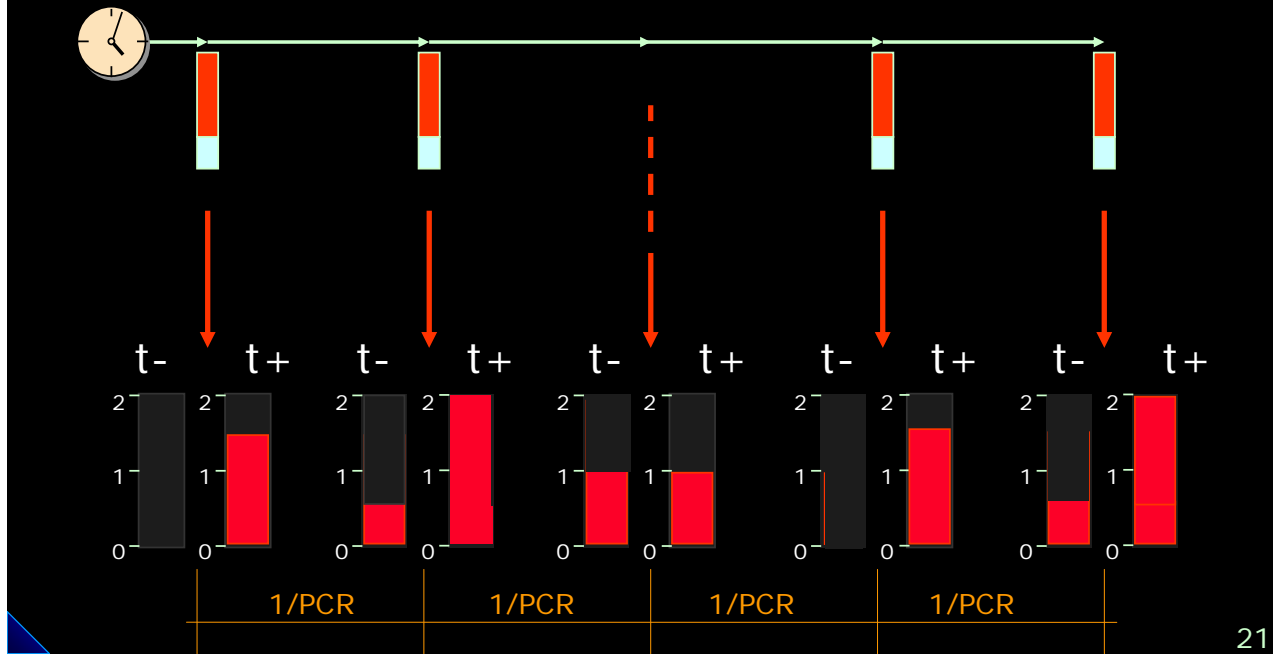
" L " e " I " determinano insieme la dimensione del Leaky Bucket.

Esempio di applicazione di GCRA con traffico costante

GCRA: $I = 1,5$; $L = 0,5$

Decremento: 1 per ogni $\Delta T = 1/PCR$

Sustainable Cell Rate = I
Burst Size = L



21

Consideriamo l'algoritmo GCRA nel caso di traffico costante.

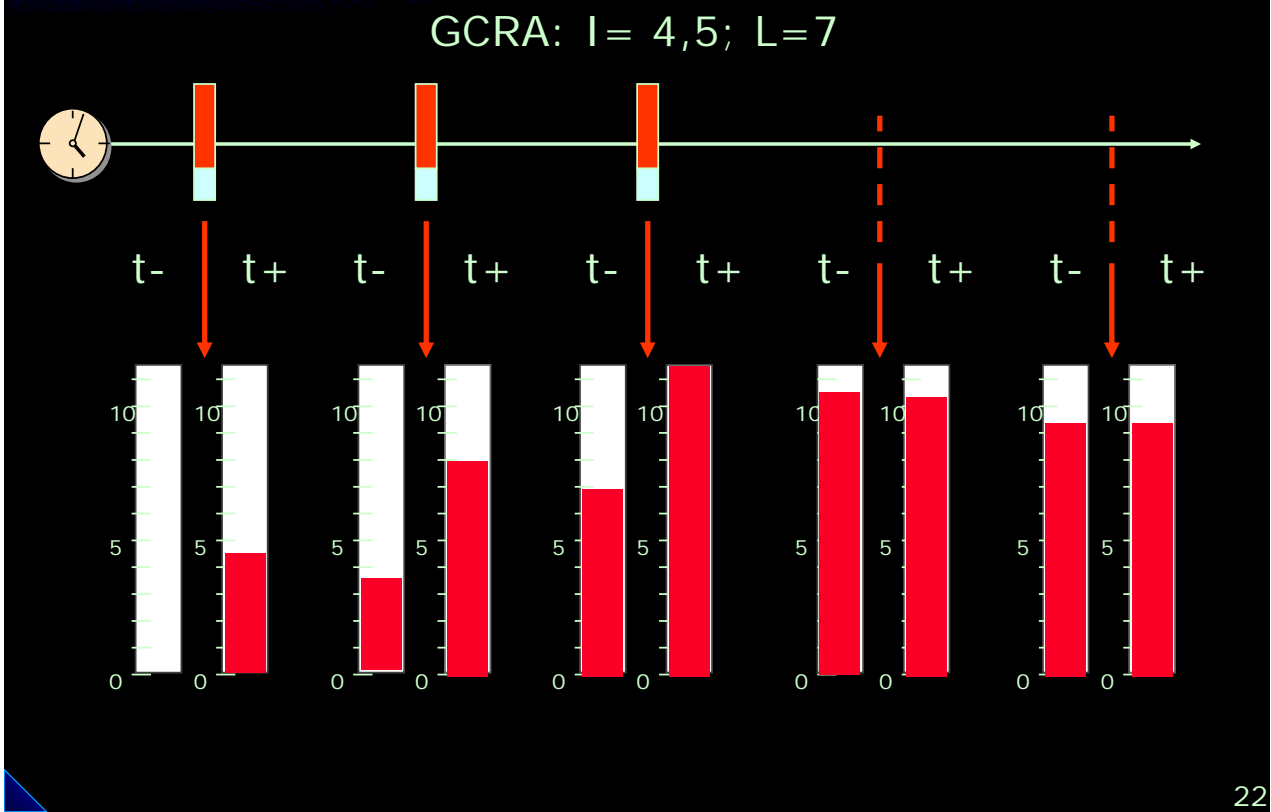
Sul diagramma viene rappresentato lo stato di riempimento del secchio un attimo prima (t^-) ed un attimo dopo (t^+) l'arrivo della cella.

Supponiamo che inizialmente il secchio sia vuoto ed una cella entri nella connessione esaminata. Il sistema di controllo mediante l'algoritmo GCRA (1,5/0,5) versa un'unità e mezzo di liquido nel secchio. Da questo momento il secchio perderà costantemente un'unità di liquido per ogni intervallo di cella (inverso del Peak Cell Rate).

Dall'esempio si capisce che con la capienza massima del secchio pari a due unità, l'arrivo di tre celle consecutive appartenenti alla stessa connessione viola il contratto.

Si nota quindi che il parametro " I " = 1,5 (3/2 celle) è il rapporto tra la massima capacità di trasferimento delle celle (Peak Cell Rate) sulla connessione e la capacità di trasferimento concessa in base al contratto stipulato (Sustainable Cell Rate).

Esempio di applicazione di GCRA con traffico a burst



22



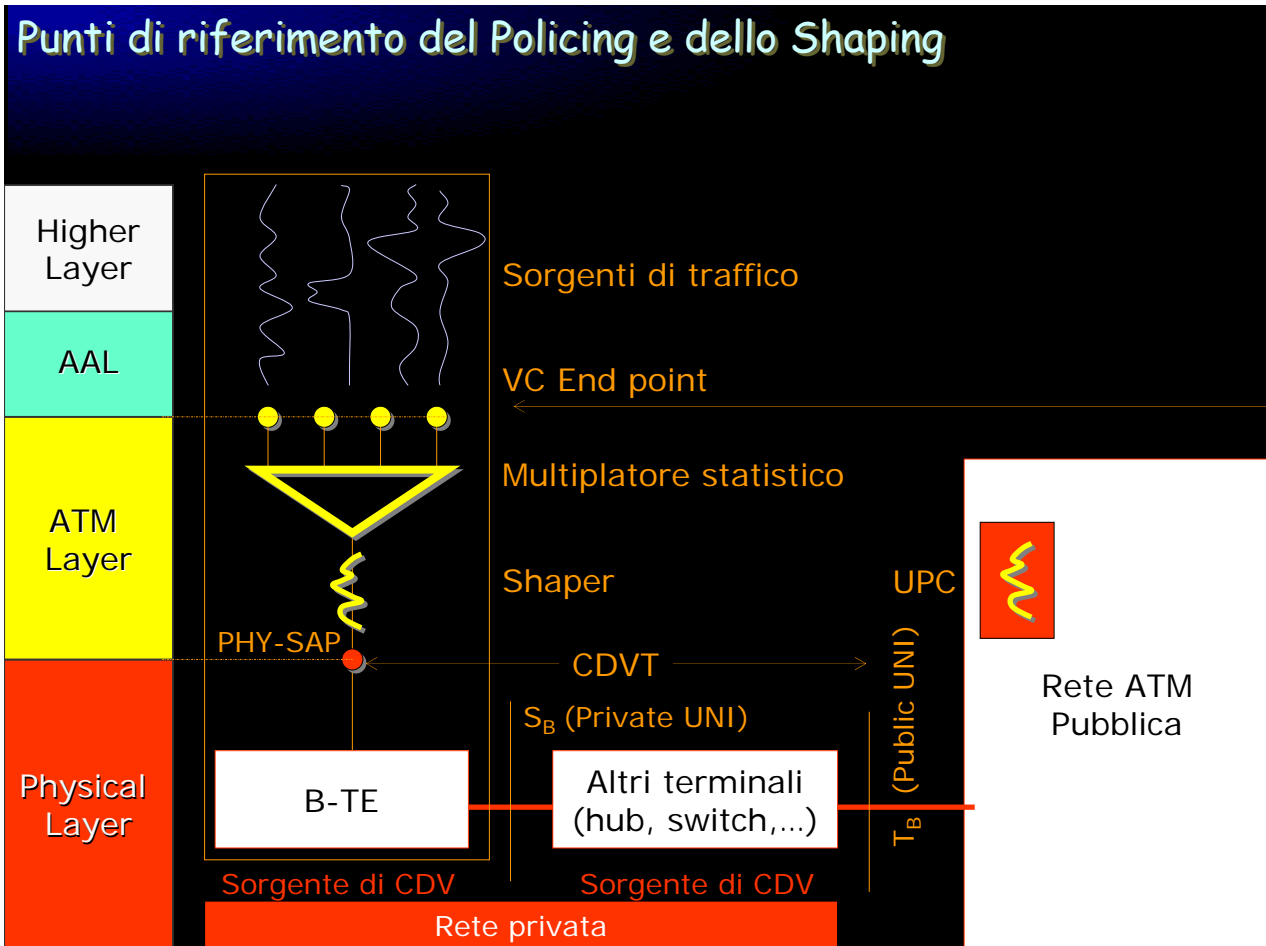
Al fine di vedere l'utilità del parametro " L ", prendiamo in considerazione un esempio di traffico a burst. Fissando " L " pari a 7 e " I " pari a 4.5, si ottiene che la capacità del secchio è pari a 11.5 unità.

Come mostra l'esempio, il secchio è pieno dopo tre celle.

Bisogna poi aspettare un tempo sufficientemente lungo affinché il secchio si svuoti completamente (11,5 tempi di cella) per trasferire un altro burst di tre celle.

Si intuisce quindi che, aumentando il valore del parametro " L ", vengono permessi burst di dimensioni maggiori. Questo riguarda specialmente le applicazioni per trasmissione dati.

Impostando i parametri " I " e " L " è possibile realizzare contratti "su misura" per ogni tipo di traffico.



L'algoritmo UPC è considerato come una forma di **policing del traffico** (sorveglianza), che consiste nel regolare un flusso dati in modo tale che le celle (o trame, o pacchetti) che superano un certo limite di prestazioni, vengano eliminate oppure marcate. E' talvolta opportuno integrare la politica del policing del traffico con una di **shaping** (sagomatura); quest'ultima è usata per smorzare l'andamento del flusso di traffico e ridurre l'addensamento delle celle, così da ottenere un'allocazione delle risorse più equa, ed una riduzione del ritardo medio. Un metodo semplice di shaping del traffico consiste nell'uso di un particolare algoritmo UPC, noto come **token bucket** (secchio di gettoni). A differenza dell'algoritmo UPC, che controlla semplicemente il traffico ed elimina o etichetta le celle non conformi al contratto, il token bucket per lo shaping del traffico opera sul flusso delle celle conformi ai parametri del contratto di traffico.

La funzione dell'UPC e del traffic shaping sono esplicitate in punti ben determinati della rete ATM. Con il modello di riferimento del terminale ATM si hanno i riferimenti logici di tali processi.

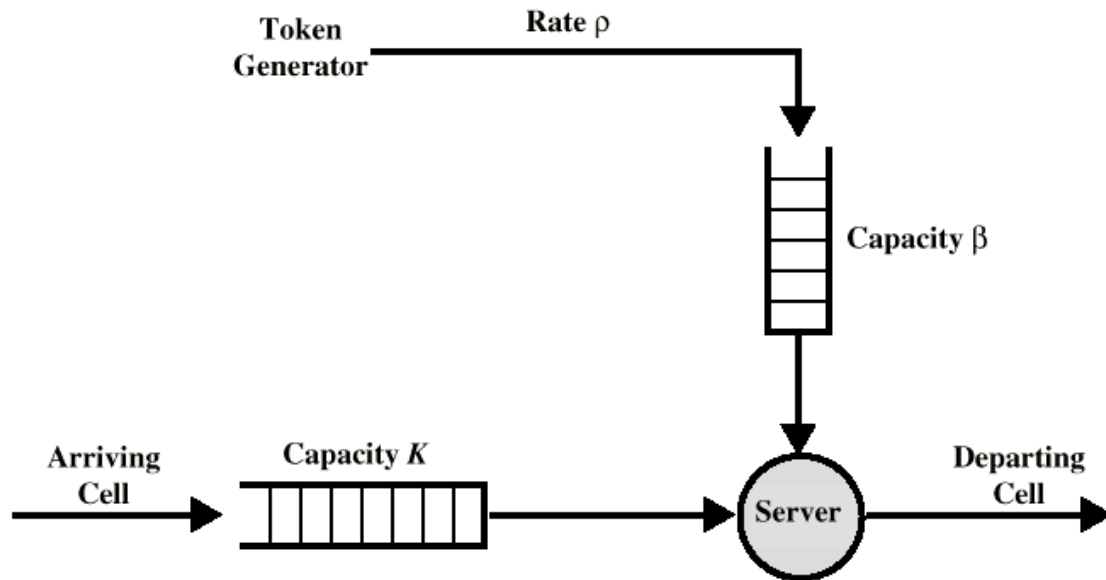
Pertanto gli elementi logici di un terminale sono a livello ATM:

Multiplicazione statistica, permette di moltiplicare su un'unica connessione fisica più sorgenti di natura statistica;

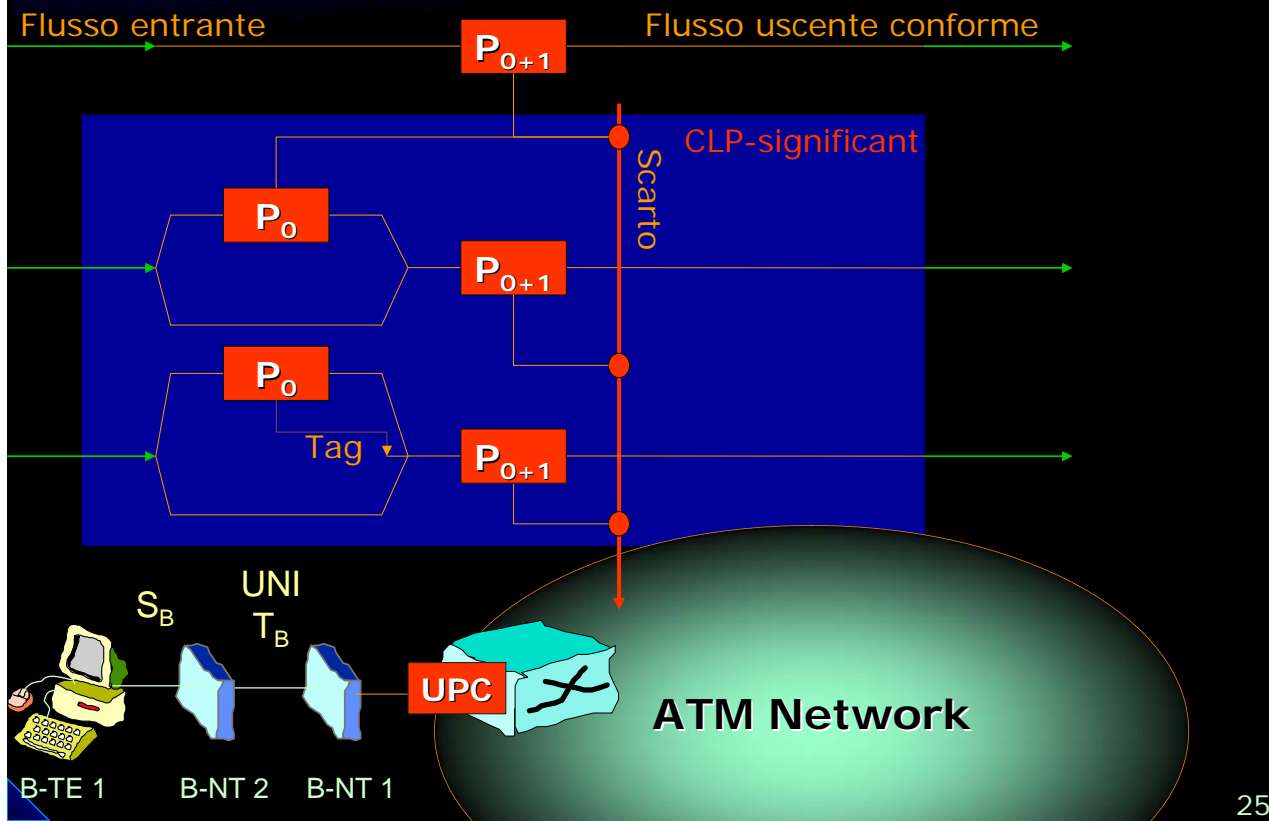
Traffic shaping, conforma il flusso moltiplicato in funzione delle caratteristiche del traffico concordate con la rete pubblica. Questa funzione può opzionalmente essere implementata anche sugli switch della rete, ma è difficilmente gestibile.

A livello fisico ci sono gli elementi che permettono l'interconnessione "hardware" alle varie interfacce sia pubbliche che private. Questi elementi possono introdurre ritardi (CDV), pertanto il parametro di tolleranza al ritardo CDVT deve essere stipulato tra il PHY-SAP e il Reference Point T_B.

Token Bucket



Policing con valutazione delle priorità



25

La verifica della conformità, definita anche policing può avvenire tenendo in considerazione della priorità delle celle all'interno di un flusso sulla connessione attivata.

Il policing può essere di tipi CLP-trasparente in cui non si dà significato al bit CLP, oppure CLP-significant in cui si differenziano all'interno del flusso le celle a bassa priorità e ad alta priorità.

Nel caso CLP-significant si può inoltre operare sul flusso ad alta priorità la diminuzione di priorità: tagging. Tale operazione che viene realizzata a fronte del mancato rispetto del contratto del traffico deve essere negoziato a priori. Per default la rete non opera il tagging e l'utente non richiede tale operazione.

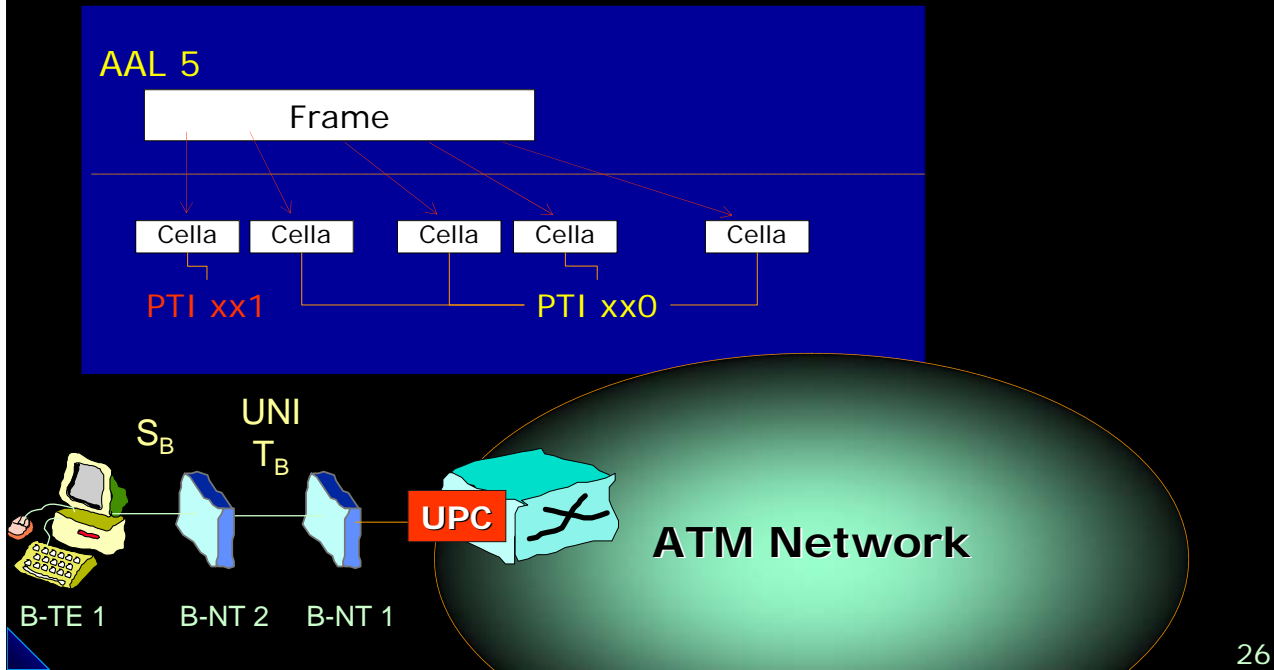
In particolare, un utente può far uso, su una connessione ATM, di due livelli di priorità, impostati da sé stesso in ogni cella, attraverso il bit CLP dell'intestazione. Quando si utilizzano i due livelli di priorità si devono specificare i parametri di traffico separatamente per entrambi i flussi di celle associati a ciascuna priorità. Generalmente questo avviene specificando un insieme di parametri di traffico associati al flusso con maggiore priorità (CLP=0) ed un insieme di parametri di traffico relativi a tutto il traffico (CLP=0 o 1) complessivo. In questo modo la rete può assegnare le risorse più efficacemente.

Ad arbitrio della rete, può essere usata la marcatura di cella in corrispondenza di celle non conformi; in questo caso, una cella non conforme può essere etichettata con CLP=1 (bassa priorità), e quindi inoltrata lo stesso. Nel caso in cui si verificasse una situazione di congestione in un certo punto della rete, le celle marcate diventano soggette ad eliminazione.

Se, in una rete, un utente ha negoziato due livelli di priorità CLP, allora la situazione diventa più complessa. Ricordando che l'utente può negoziare un contratto di traffico per il traffico con maggiore priorità (CLP=0), ed un altro per il traffico globale (CLP=0 o 1), le regole che vengono applicate sono le seguenti:

1. Una cella con CLP=0, che è conforme al contratto di traffico previsto per CLP=0, è inoltrata normalmente;
2. Una cella con CLP=0, che non è conforme al traffico previsto per CLP=0, ma è conforme al traffico per CLP=0 o 1, viene passata alla priorità inferiore e quindi inoltrata;
3. Una cella con CLP=0, che non è conforme al traffico previsto per CLP=0 e nemmeno a quello per CLP=0 o 1, viene scartata;
4. Una cella con CLP=1, che è conforme al contratto di traffico previsto per CLP=0 o 1, è inoltrata normalmente;
5. Una cella con CLP=1, che non è conforme al contratto di traffico previsto per CLP=0 o 1, viene scartata.

Frame discard



L'eliminazione selettiva di celle entra in gioco quando la rete, in un certo punto successivo a quello in cui agisce la funzione UPC, scarta le celle con CLP=1. L'obiettivo è quello di eliminare, durante una congestione, le celle con priorità più bassa per salvaguardare le prestazioni delle celle a maggiore priorità. A tal proposito, si noti che la rete non ha la possibilità di distinguere le celle che sono state etichettate come a bassa priorità dalla sorgente, da quelle che sono state invece etichettate dalla funzione UPC.

In questa accezione la trama è da intendersi come quella organizzazione di bit (AAL-PDU) che viene segmentata in tante celle dai protocolli di adattamento. In particolare il protocollo AAL 5 utilizza il terzo bit del Payload Type Identifier per segnalare l'appartenenza delle celle alla stessa AAL-PDU (le modalità sono state viste precedentemente). Generalmente i nodi della rete non processano tale bit, ma potrebbe diventare utile in situazione di congestione del traffico.

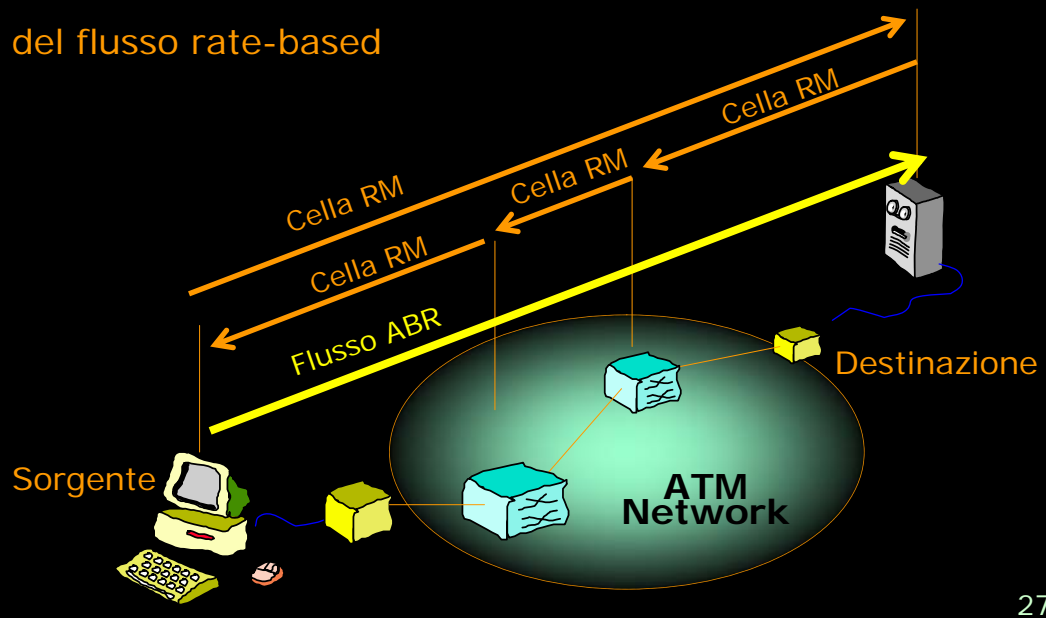
Facciamo un esempio di interconnessione tra reti LAN. La trama Ethernet ha una lunghezza variabile fino a 1500 byte, ciò significa che il protocollo AAL con la funzione di SAR "genererà" circa 300 celle per trama. Nel caso di scarto di una cella singola i protocolli applicativi sulle LAN interconnesse richiederanno la ritrasmissione dell'intera trama, e cioè di altre 300 celle!

Sarebbe molto efficace, nel caso in cui si debba scartare una cella, scartare anche tutte quelle appartenenti alla stessa trama. Ciò libererebbe risorse in modo ottimizzato.

Controllo del flusso ABR

Gli obiettivi dell'ABR sono:

- minimizzare la quantità di traffico perso;
- ottimizzare l'occupazione delle risorse disponibili;
- Controllo del flusso rate-based



27



La categoria di servizio ABR richiede meccanismi di controllo del flusso mediante un flusso di celle di servizio che formano dei loop. Questo serve per adattare istantaneamente la velocità di trasferimento delle celle. L'ABR non ha pertanto meccanismi di policing con parametri stabiliti in fase di attivazione della connessione, ma concede istantaneamente le risorse disponibili.

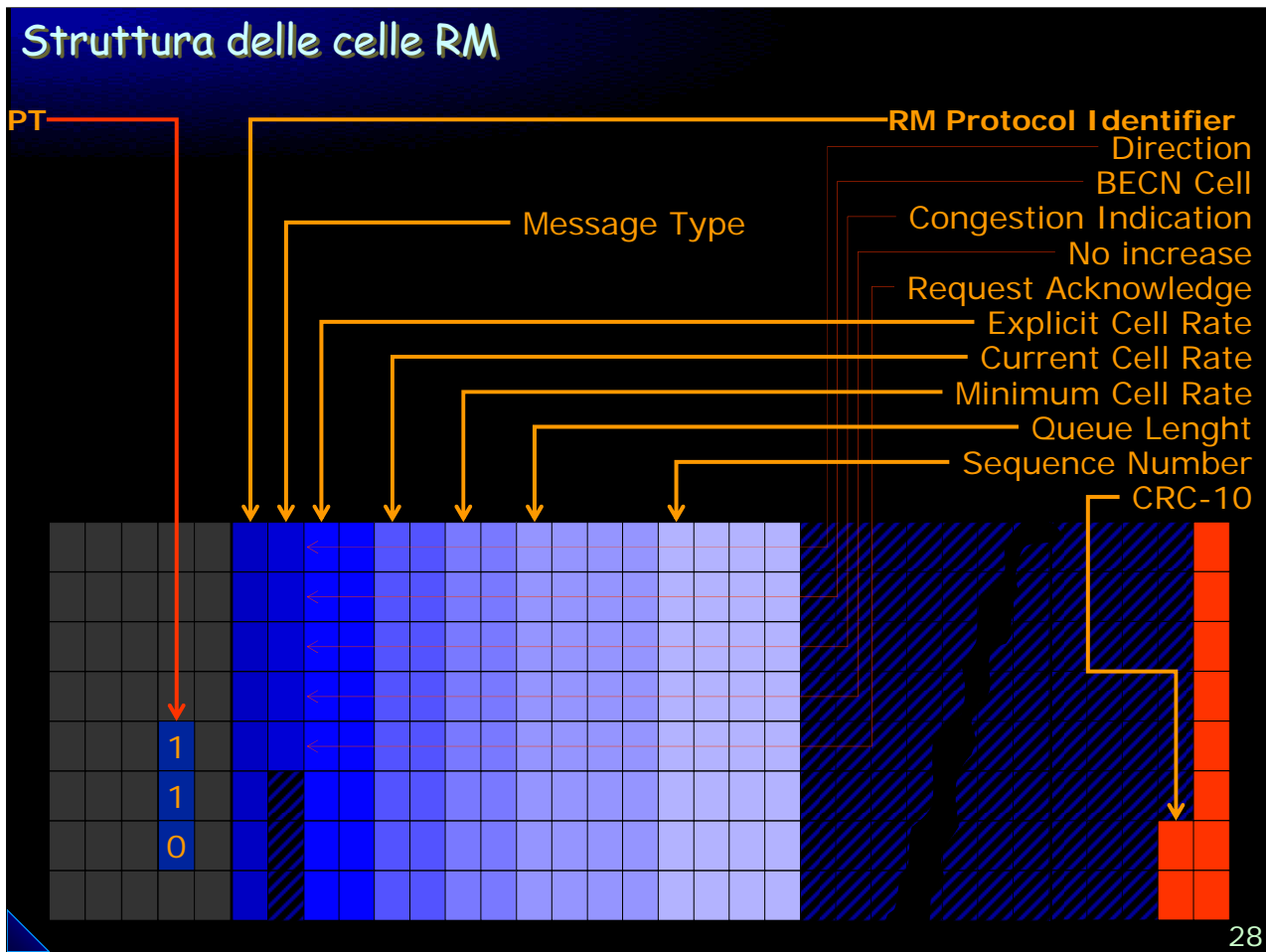
Gli obiettivi sono:

minimizzare la quantità di traffico perso;

ottimizzare l'occupazione delle risorse disponibili;

Il controllo del flusso è di tipo rate-based che consiste nelle seguenti operazioni:

- la stazione sorgente invia una cella RM (Resource Management) verso la stazione di destinazione con la richiesta della banda necessaria;
- la rete inoltra la richiesta verso la destinazione, la quale conferma o modifica la richiesta di banda in funzione alle proprie disponibilità;
- la stazione di destinazione manda alla stazione sorgente una cella RM con tali indicazioni;
- durante il tragitto di ritorno i nodi interessati alla connessione partecipano al feedback indicando le proprie disponibilità sulla cella RM di feedback;
- alla ricezione della cella RM la stazione sorgente trasmette alla velocità minore indicata



28

Descrizione dei campi delle celle RM

Header: The first five bytes of an RM-cell are the standard ATM header with PTI=110 (binary) for a VCC, and additionally VCI=6 for a VPC. The CLP bit is 0 if the RM-cell is in-rate and 1 if it is out-of-rate.

Payload

ID: The protocol ID identifies the service using the RM-cell. The ITU has assigned protocol ID = 1 to ABR service.

Message Type Field

DIR: The DIR bit indicates which direction of data flow is associated with the RM-cell. A forward RM-cell, indicated by DIR=0, is associated with data cells flowing in the same direction. A backward RM-cell, indicated by DIR=1, is associated with data cells flowing in the opposite direction. DIR is changed from 0 to 1 when an RM-cell is turned around at a destination.

BN: The BN bit indicates whether the RM-cell is a Backward Explicit Congestion Notification (BECN) cell (i.e., non-source generated) or not. BN=0 indicates a source generated RM-cell while BN=1 indicates a BECN RM-cell generated by a destination or a switch.

CI: The CI (congestion indication) bit allows a network element to indicate that there is congestion in the network. When a source receives a backward RM-cell with CI=1 it decreases its ACR. When turning around a forward RM-cell, a destination will set CI=1 to indicate that the previous received data cell had the EFCI state set.

NI: The NI (no increase) bit is used to prevent a source from increasing its ACR. In contrast to CI=1, NI=1 does not require any decrease. A network element might set NI to 1 to indicate impending congestion. Normally, a source will initialize NI to 0 so that it might be allowed to increase its ACR, but it can indicate that it does not need a higher ACR by initializing NI to 1.

RA: The RA bit is not used for ATM Forum ABR.

ER: The ER (Explicit Rate) field is used to limit the source ACR to a specific value. For each RM-cell ER is set by the source to a requested rate (such as PCR). It may be subsequently reduced by any network element in the path to a value that the element can sustain.

CCR: The CCR field is set by the source to its current ACR. It may be useful to network elements in computing a value to place in ER. For BECN cells, CCR=0.

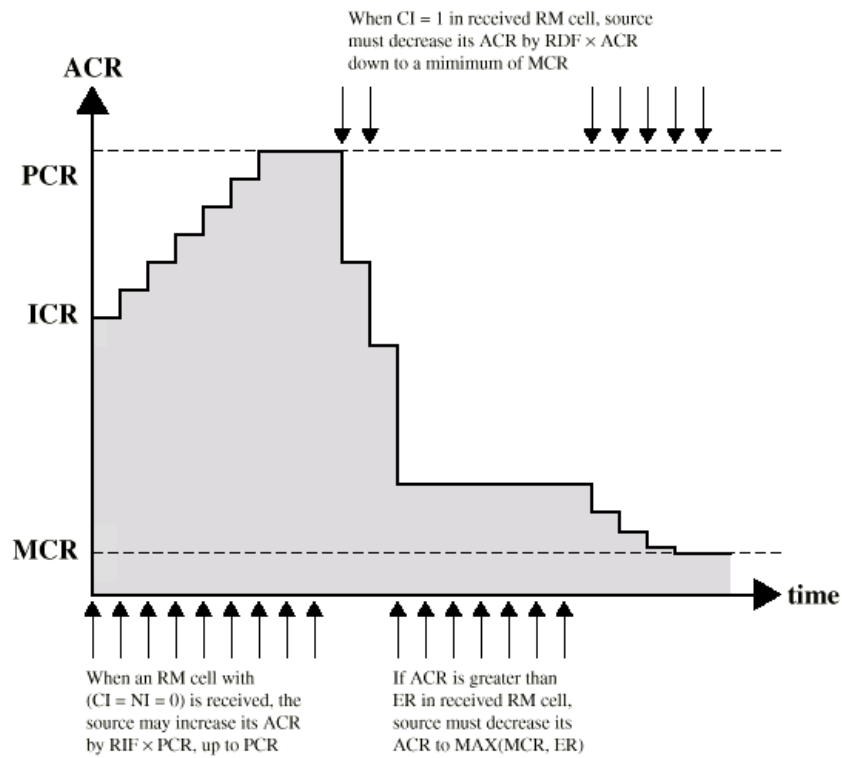
MCR: The MCR field carries the connection's Minimum Cell Rate. It may be useful to network elements in allocating bandwidth among connections. For BECN cells, MCR=0.

QL: The QL field is not used for ATM Forum ABR.

SN: The SN field is not used for ATM Forum ABR.

CRC-10: The RM CRC is the same CRC used for all OAM cells. It is computed as the remainder of the division (modulo 2) by the generator polynomial of the product of x^{10} and the content of the RM-cell payload excluding the CRC field (374 bits). Each bit of this payload is considered as a coefficient (modulo 2) of a polynomial of degree 373 using the first bit as the coefficient of the highest order term. The CRC-10 generating polynomial is: $1 + x^4 + x^5 + x^9 + x^{10}$. The result of the CRC calculation is placed with the least significant bit right justified in the CRC field.

Variations in Allowed Cell Rate



29

Cell Flow

