浙江大学计算机学院
数字媒体与网络技术

Digital Asset Management
数字媒体资源管理

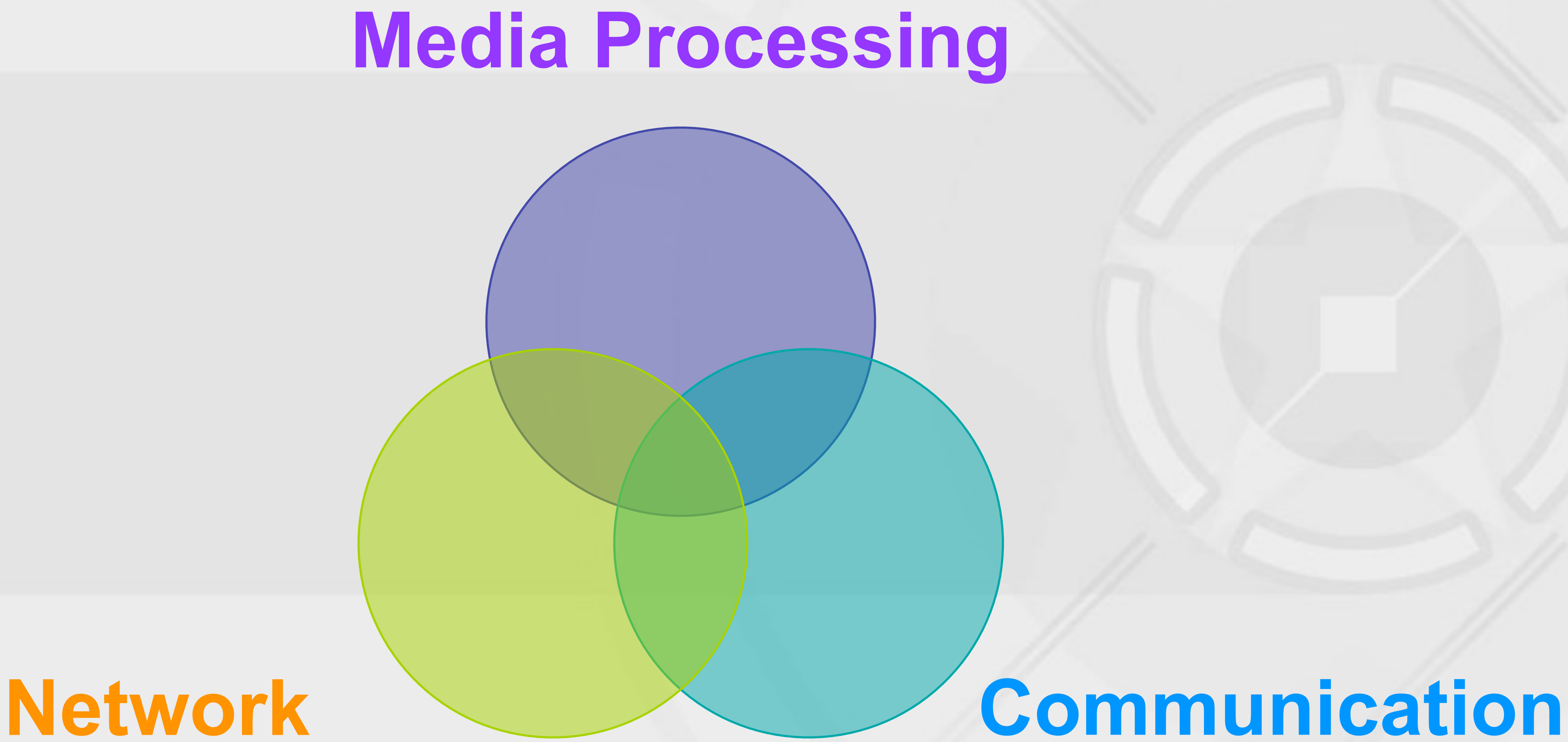
5. Streaming multimedia

任课老师：张宏鑫
2015-10-27

Keys of Streaming Media

- Algorithms (**)
- Standards (*****)
- Complete End-to-End systems (***)
- Research Frontiers(*)

Ubiquitous Multimedia



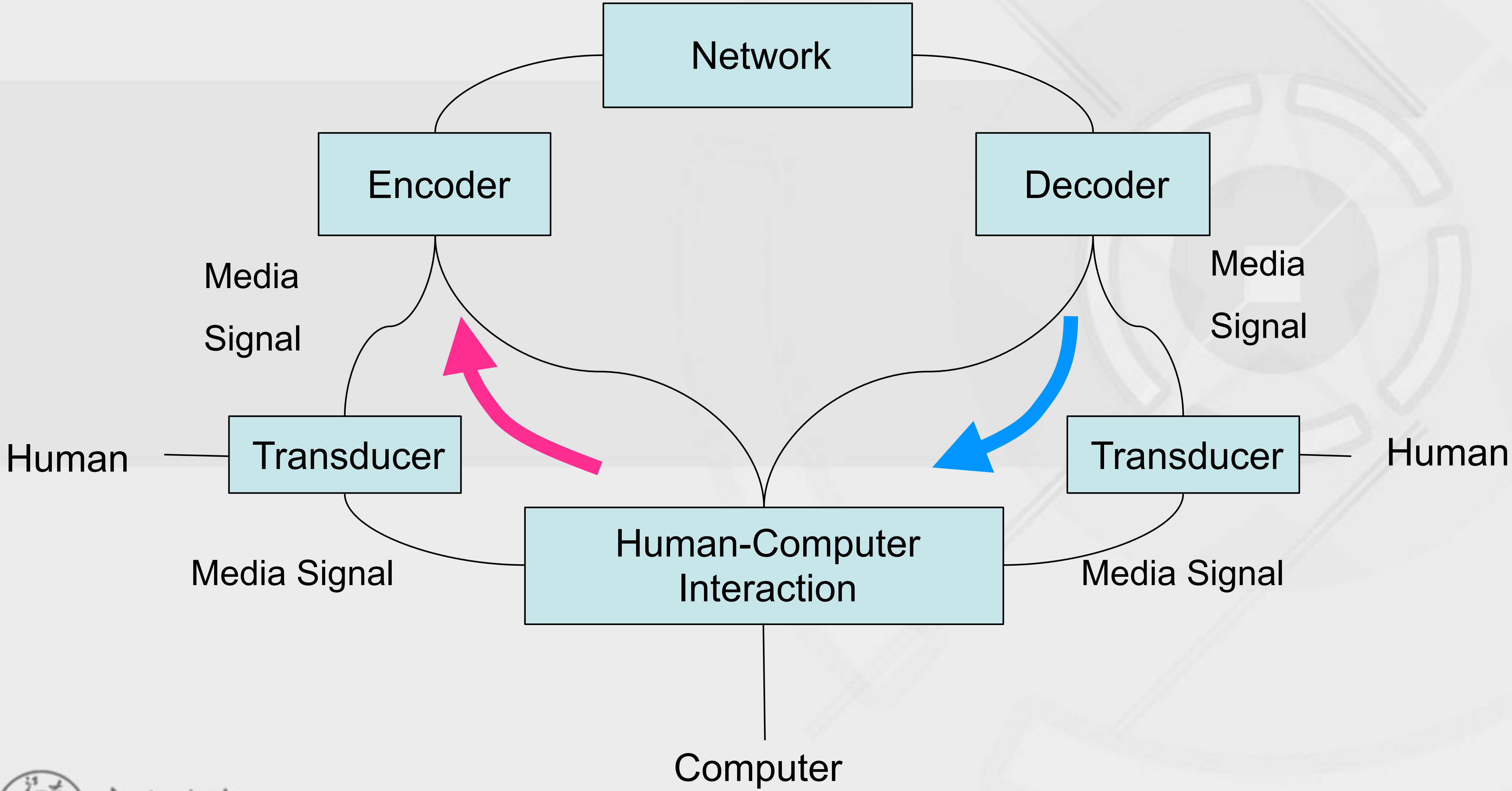
The dimensions of multimedia

- Audio, Video
- Images
- Graphics
- Animation
- Text, Data files, Speech, Handwriting

Key Dimensions of Communication Tech

- Antenna Diversity and Space-Time Processing
- Channel Equalization
- Bitstream Interleaving
- Multiple Access and Interference Cancellation
- Modulation and Error Protection

Multimedia Communications System



Key Dimensions of Networking

- Media Access Protocols
- Error Control: ARQ
- Admission Control, Scheduling
- Routing, Multicasting
- Mobility Management and User Tracking
- Protocols for ATM, IP and Hybrid Networks

Multimedia in networks (1): Characteristics

Fundamental characteristics:

- Typically **sensitive delay**.
- But **loss tolerant**: infrequent losses cause minor glitches that can be concealed.
- **Antithesis of data** (programs, banking info, etc.), which are loss intolerant but delay tolerant.
- Multimedia is also called "**continuous media**"

Classes of MM applications:

- Streaming stored audio and video
- Streaming live audio and video
- Real-time interactive video

Multimedia in networks (2): Applications

Streaming stored MM

- Clients request audio/video files from servers and pipeline reception over the network and display
- Interactive: user can control operation (similar to VCR: pause, resume, fast forward, rewind, etc.)
- Delay: from client request until display start can be 1 to 10 seconds

Unidirectional Real-Time:

- similar to existing TV and radio stations, but delivery over the Internet
- Non-interactive, just listen/view

Interactive Real-Time :

- Phone or video conference
- More stringent delay requirement than Streaming & Unidirectional because of real-time nature
- Video: < 150 msec acceptable
- Audio: < 150 msec good, <400 msec acceptable

Multimedia in networks (3): Challenges

- **TCP/UDP/IP** suite provides best-effort, no guarantees on delay or delay variation.
  - Streaming apps with initial delay of 5-10 seconds are now commonplace, but performance deteriorates if links are congested (transoceanic)
  - Real-Time Interactive apps have rigid requirements for **packet delay and jitter**.
  - **Jitter** is the variability of packet delays within the same packet stream.

- Design for multimedia apps would be easier if there were some 1st and 2nd class services (**QoS**).
  - But in the public Internet, all packets receive equal service.
  - Packets containing real-time interactive audio and video stand in line, like everyone else.
- There have been, and continue to be, efforts to provide differentiated service.

Multimedia in networks (4): making the best of best effort

To mitigate impact of "best-effort" Internet, we can:

- Use UDP to avoid TCP and its slow-start phase…
- Buffer content at client and control playback to remedy jitter
- We can timestamp packets, so that receiver knows when the packets should be played back.
- Adapt compression level to available bandwidth
- We can send redundant packets to mitigate the effects of packet loss.

We will discuss all these "tricks"

How should the Internet evolve to better support multimedia?

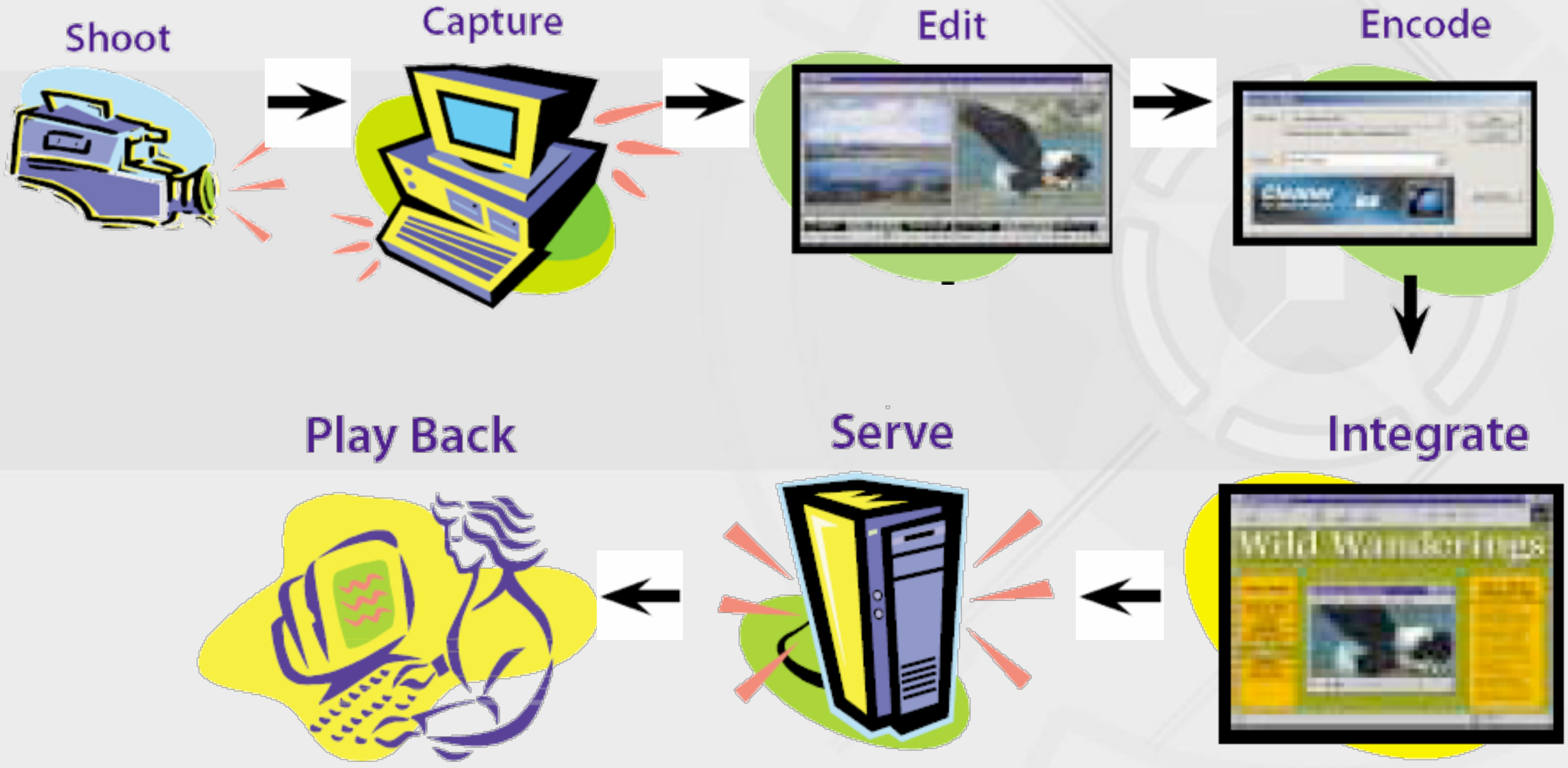
Integrated services philosophy:

- Change Internet **protocols** so that applications can reserve end-to-end bandwidth
  - Need to deploy protocol that reserves bandwidth
  - Must modify scheduling policies in routers to honor reservations
  - Application must provide the network with a description of its traffic, and must further abide to this description.
- Requires new, complex software in hosts & routers

Differentiated services philosophy:

- Fewer changes to Internet infrastructure, yet provide 1st and 2nd class service.
- Datagrams are marked.
- User pays more to send/receive 1st class packets.
- ISPs pay more to backbones to send/receive 1st class packets.

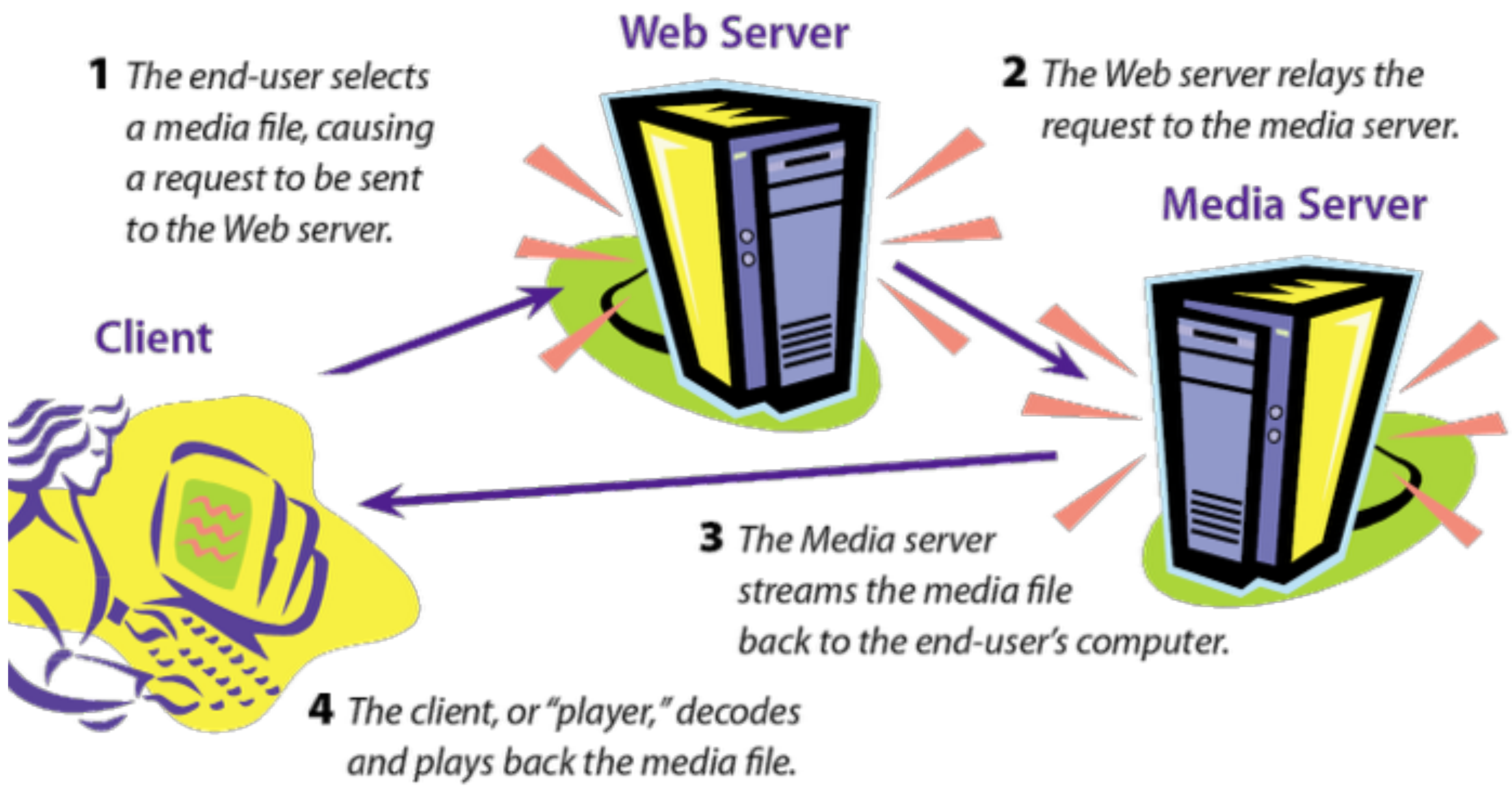
Workflow of Streaming Media



The primary characteristics of “streaming media”

- Three primary characteristics combine to define streaming media
  - –Streaming media technology enables real-time or on-demand access to multimedia content via the Internet or an intranet.
  - –Streaming media is transmitted by a media server application, and is processed and played back by a client player application, as it is received.
  - –A streamed file is received, processed, and played simultaneously and immediately, leaving behind no residual copy of the content on the receiving device.

HOW DOES STREAMING WORK?

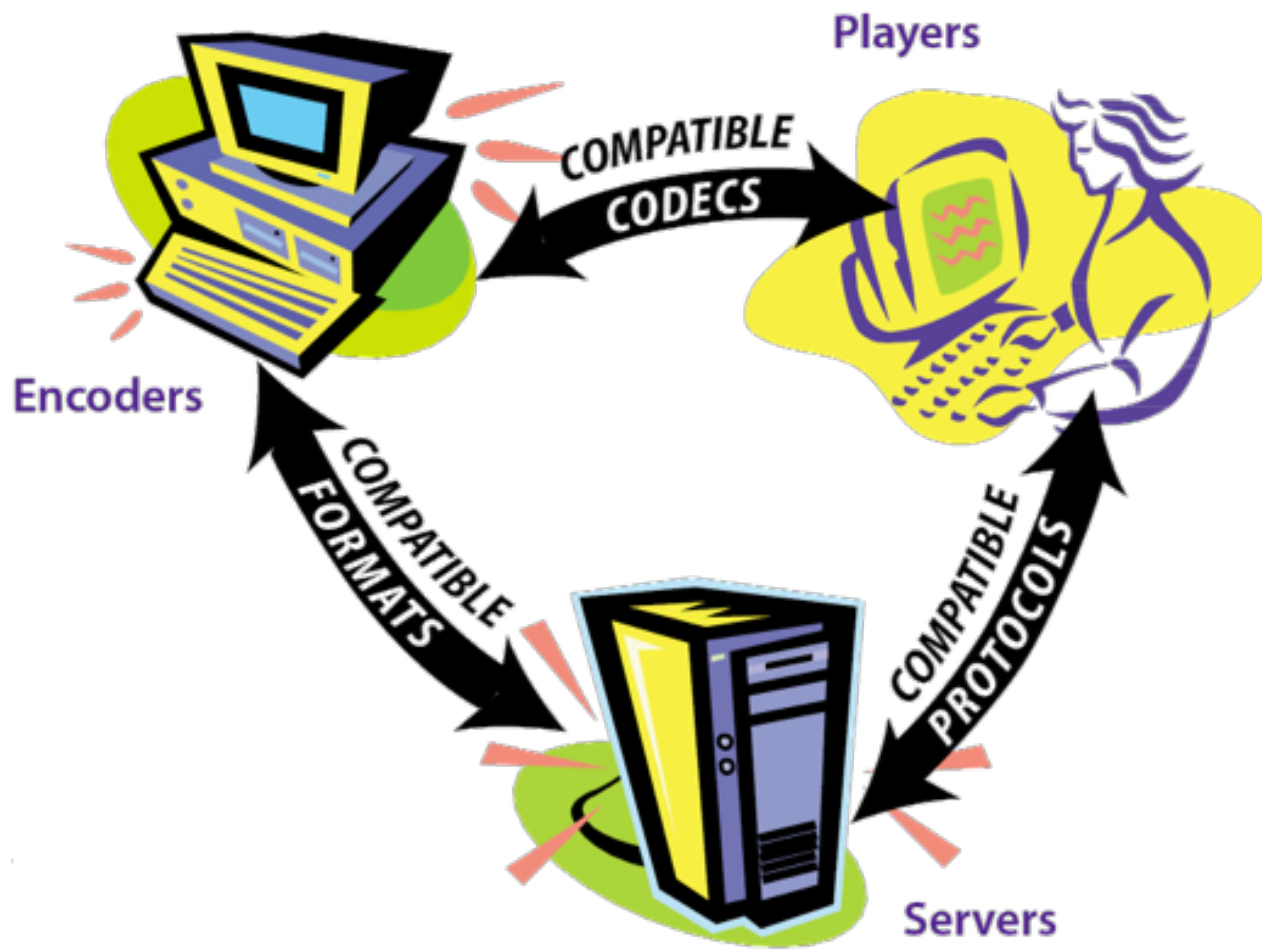


WHERE DO STREAMS COME FROM?

- Streaming media architectures.
  - Streaming media architectures are comprised of
    - encoding and transmission methods,
    - server software, and
    - players (client software).
  - The three most popular streaming media architectures
    - RealMedia,
    - Windows Media, and
    - QuickTime.

It is all interrelated

- In a streaming architecture, everything must be compatible.



Streaming media formats

Architecture	Native Formats	Streaming Media File Extensions
QuickTime	QuickTime Format	.mov (sometimes .qt or .qti)
RealMedia	RealMedia Format	.rm
Windows Media	Advanced Streaming Format or Windows Media Video/Audio	.asf, .wmv, .wma

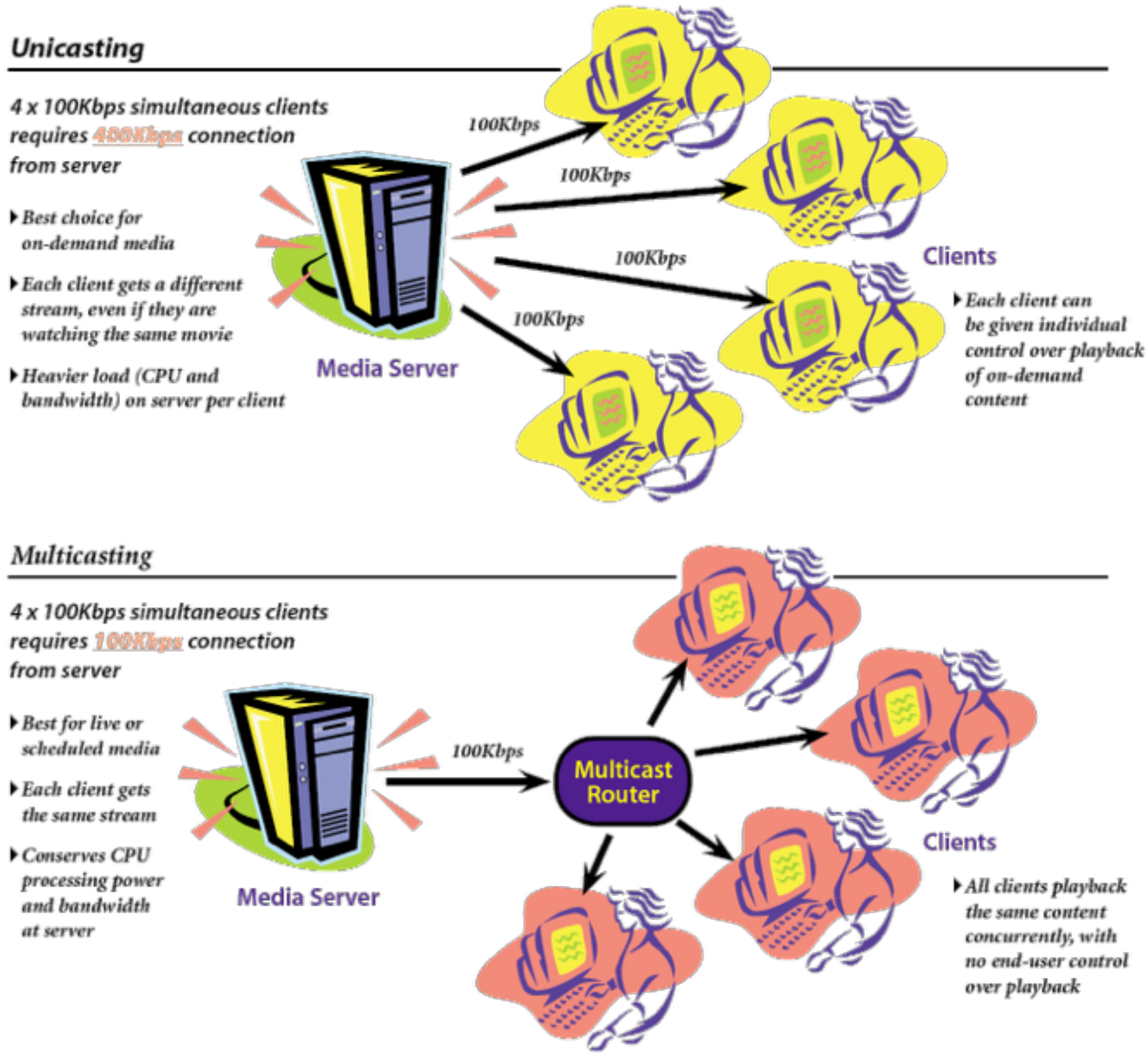
- **MPEG standard**
  - –Windows Media Video v1 is a derivative of the MPEG-4 codec, which has been renamed to avoid confusion.
  - –QuickTime 5 is the first full implementation of MPEG-4 for streaming media.

Streaming ...

- Progressive streaming transport (PST)
  - –use HTTP
  - –no jump

- Real-time streaming transport
  - –Real server (Real-time streaming protocol, RTSP)
  - –Windows Media server (M$ media server, MMS)
  - –Quicktime server

Unicast and Multicast

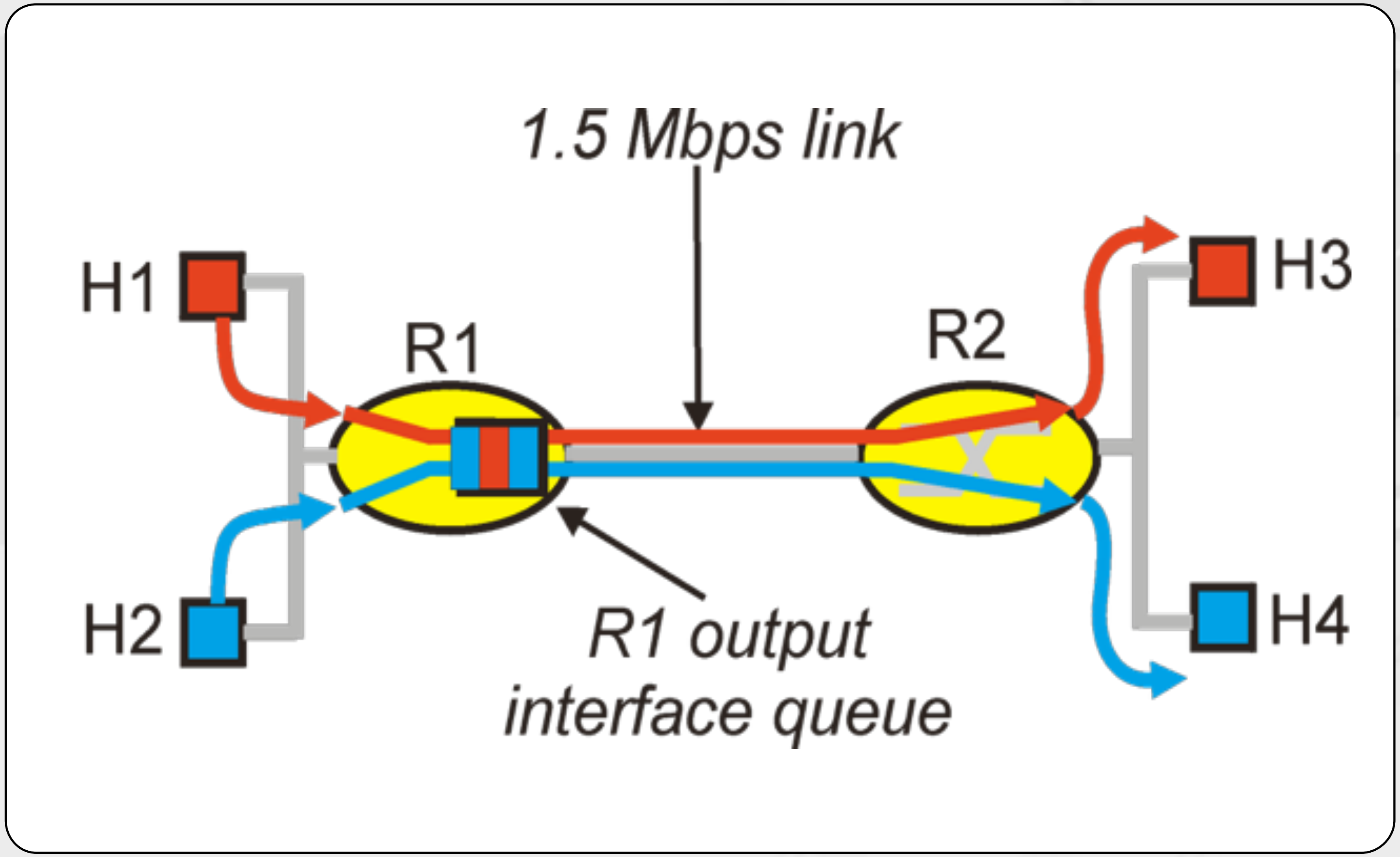


QoS: Quality of Service

- A defined measure of performance in a data communications system

- resource reservation control mechanisms
  - –make the actual determination of which packets have priority
    1. provide different priority to different users or data flows, or
    2. guarantee a certain level of performance to a data flow in accordance with requests from the application program

QoS: a simple example

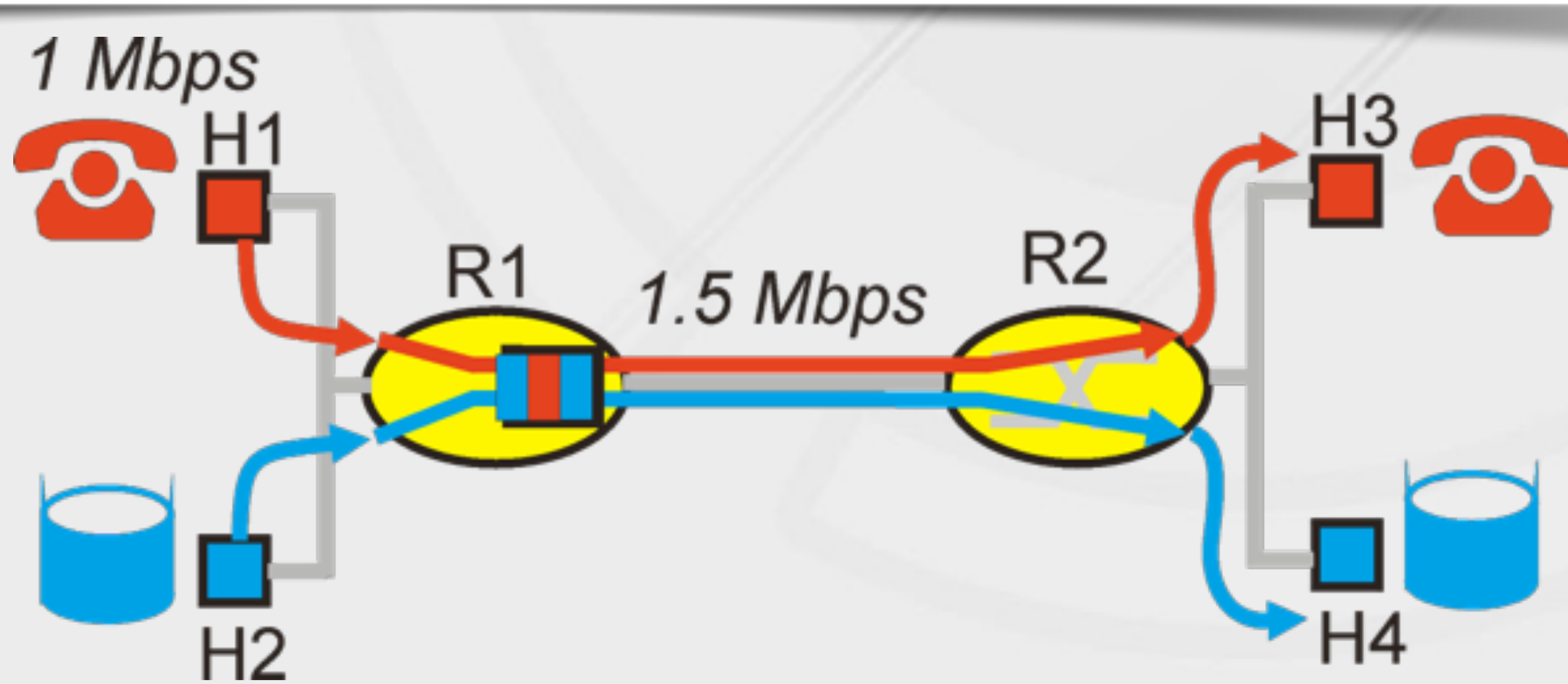


Improving QOS in IP Networks

Principles for QOS Guarantees

- **Consider:**
  - –a phone application at 1Mbps and
  - –an FTP application sharing a 1.5 Mbps link.
    - bursts of FTP can congest the router and cause audio packets to be dropped.
    - want to give priority to audio over FTP

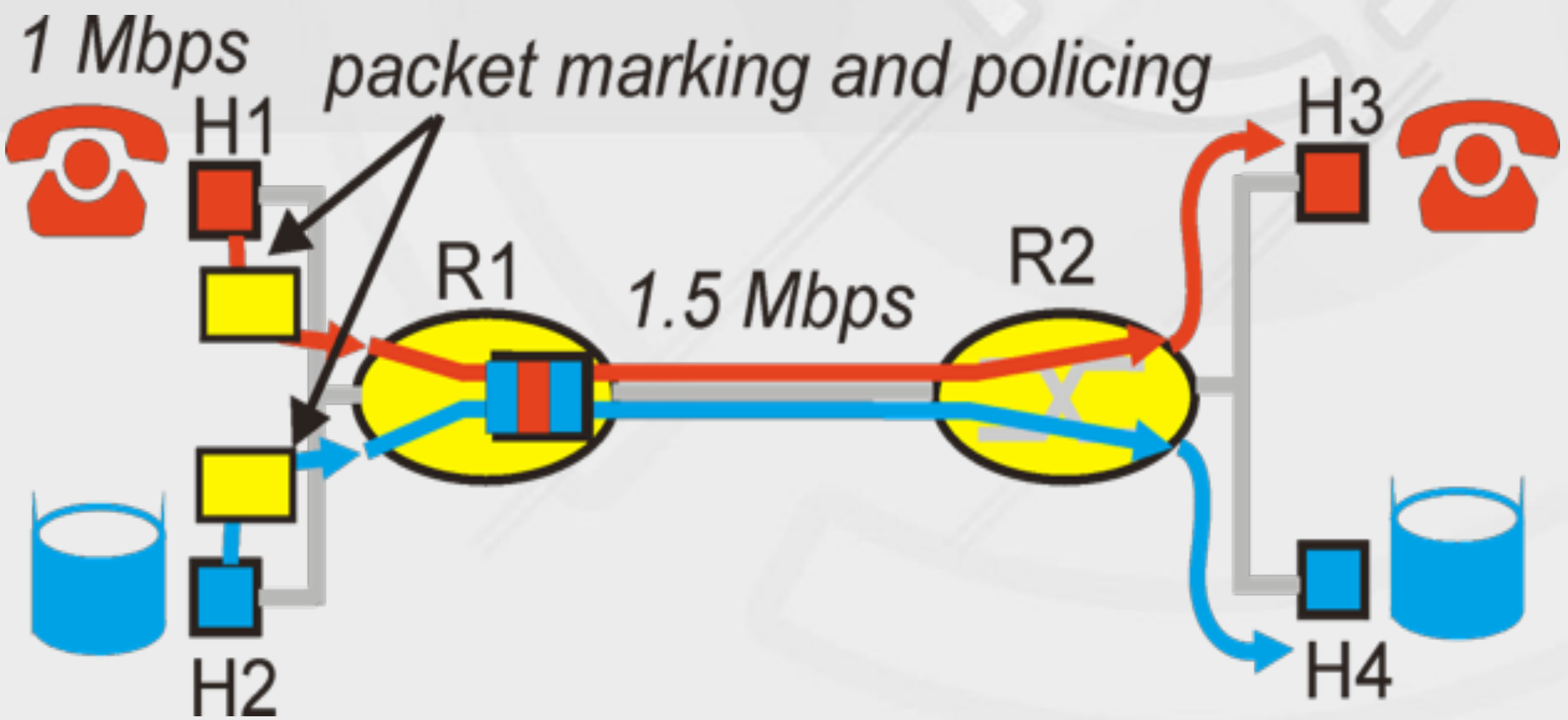
- **PRINCIPLE 1: Marking of packets is needed for router to distinguish between different classes; and new router policy to treat packets accordingly**



浙江大学计算机学院
数字媒体与网络技术

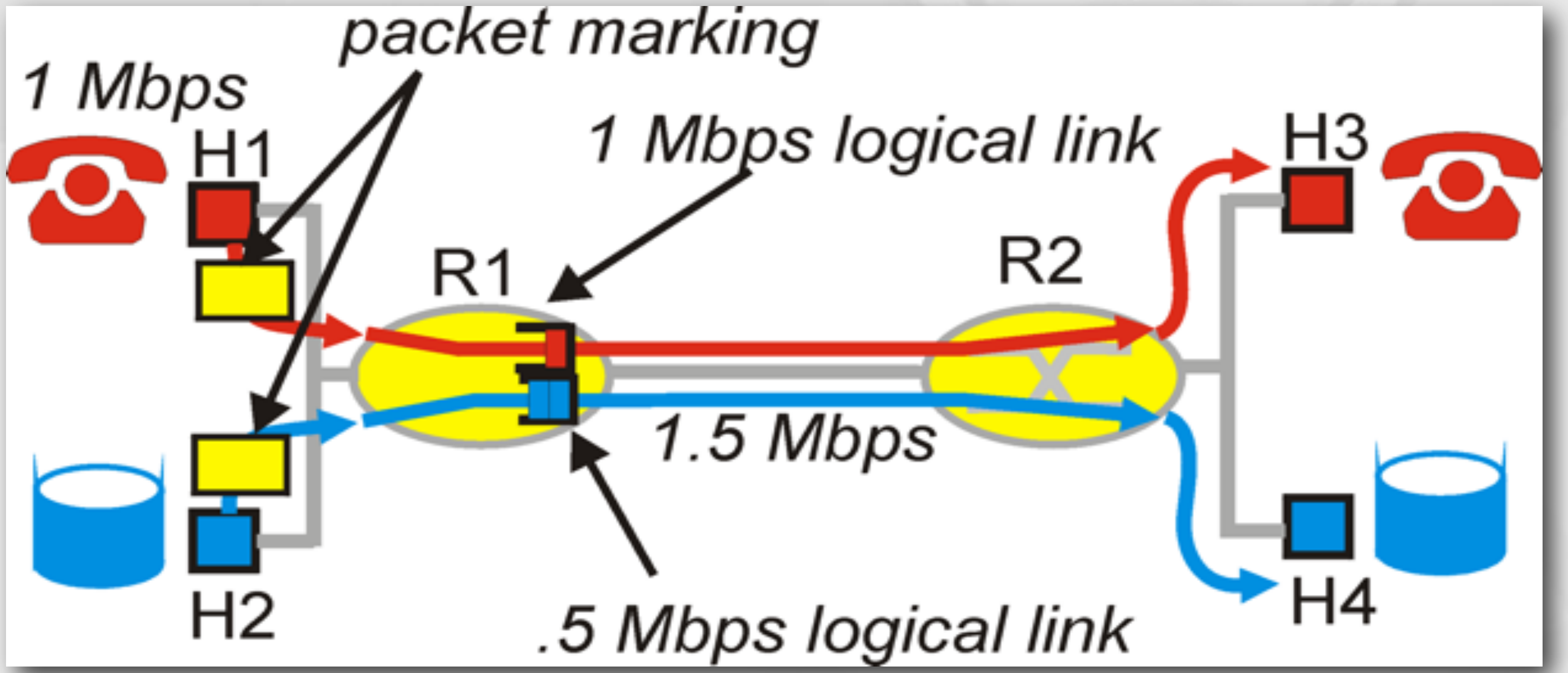
Principles for QOS Guarantees (more)

- Applications misbehave (audio sends packets at a rate higher than 1Mbps assumed above);
- **PRINCIPLE 2: provide protection (isolation) for one class from other classes**
- Require Policing Mechanisms to ensure sources adhere to bandwidth requirements; Marking and Policing need to be done at the edges:



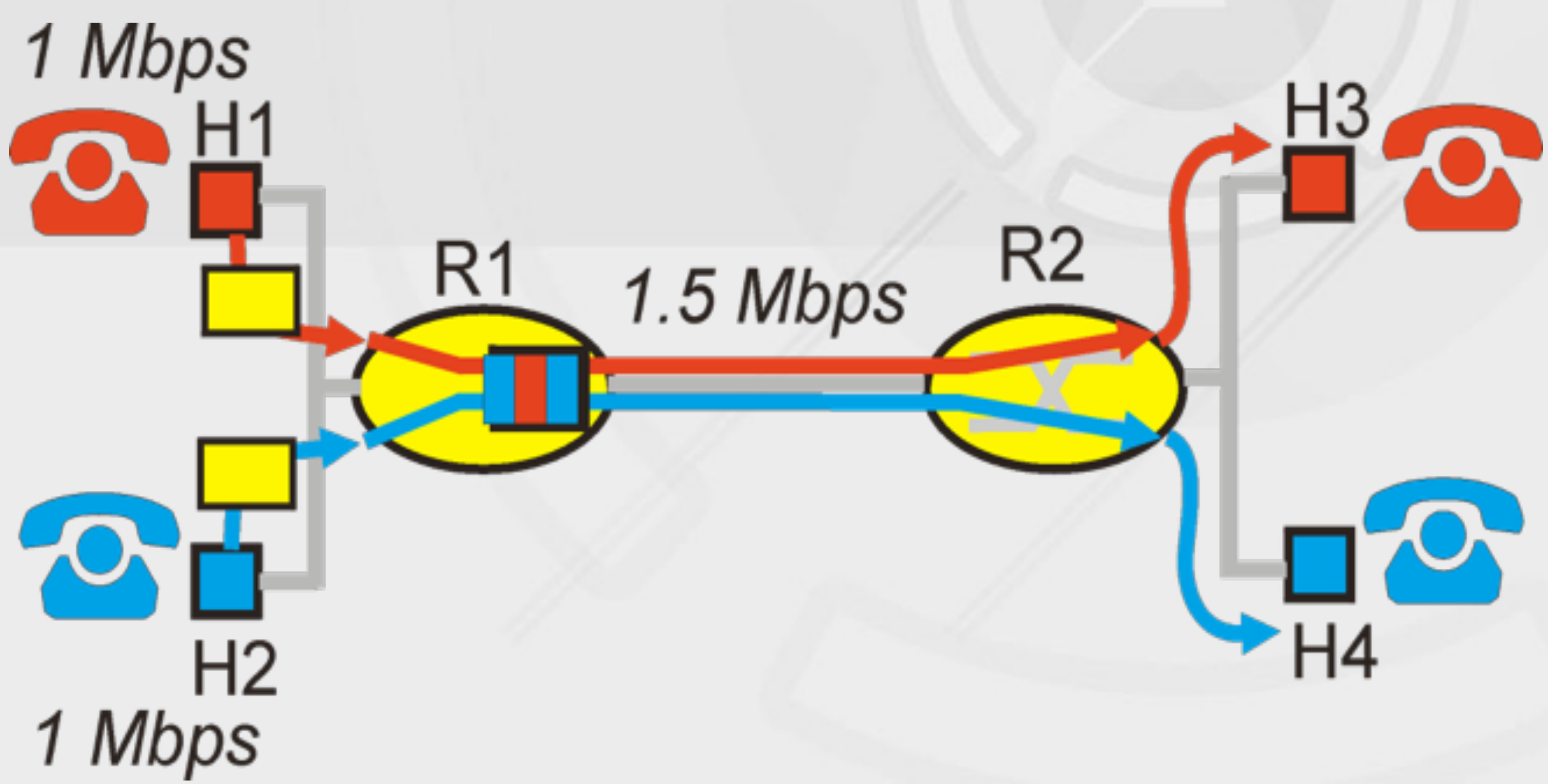
Principles for QOS Guarantees (more)

- Alternative to Marking and Policing: allocate a set portion of bandwidth to each application flow; can lead to inefficient use of bandwidth if one of the flows does not use its allocation
- **PRINCIPLE 3: While providing isolation, it is desirable to use resources as efficiently as possible**



Principles for QOS Guarantees (more)

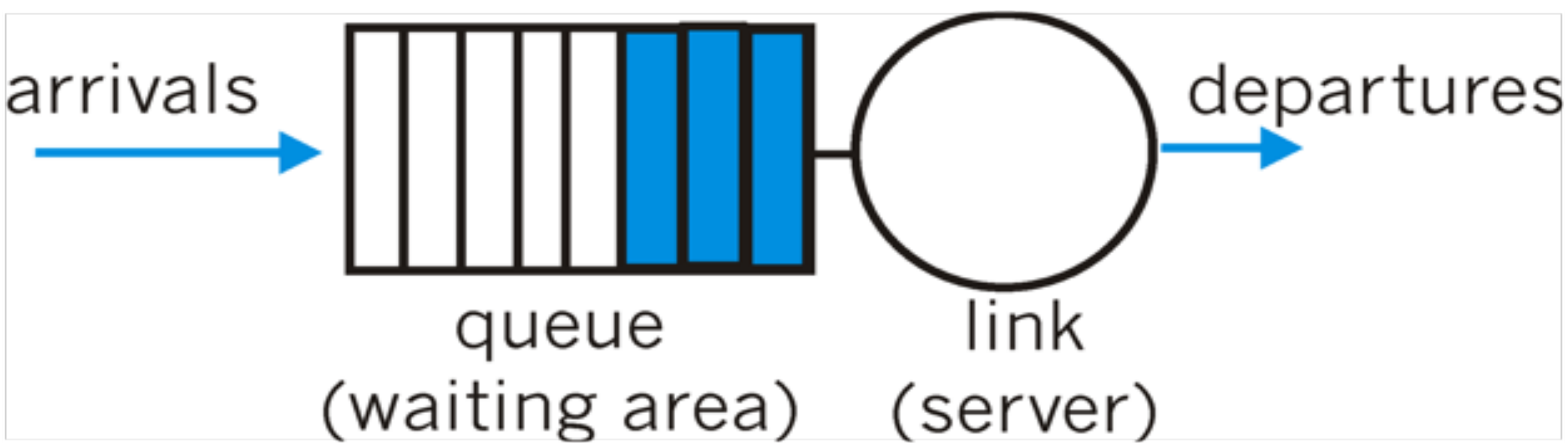
- Cannot support traffic beyond link capacity
- **PRINCIPLE 4: Need a Call Admission Process; application flow declares its needs, network may block call if it cannot satisfy the needs**



Scheduling And Policing Mechanisms

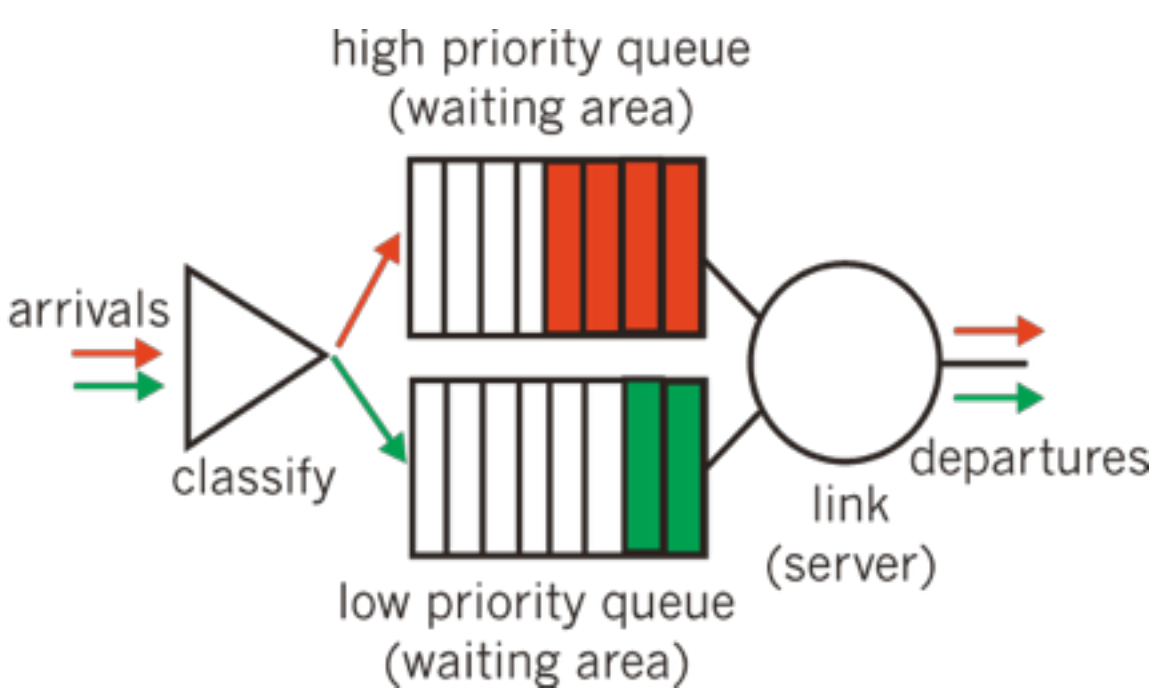
- **Scheduling:** choosing the next packet for transmission on a link can be done following a number of policies;

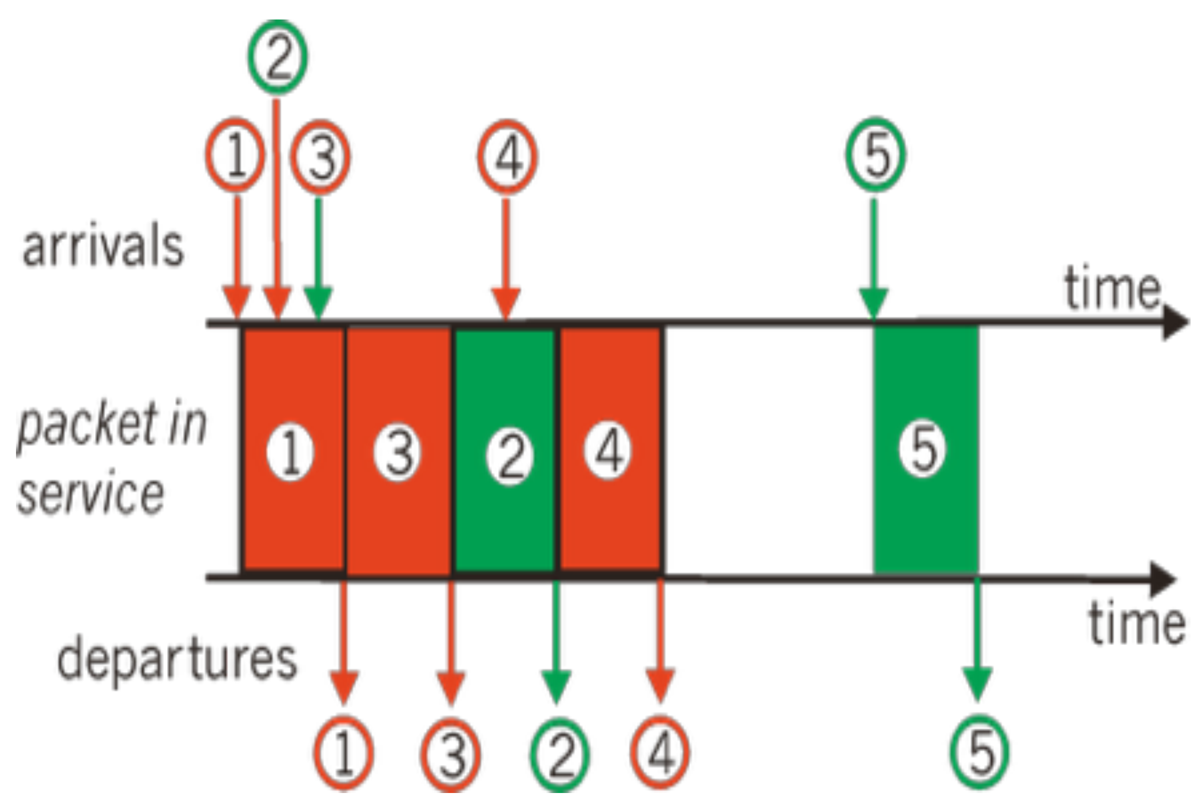
- **FIFO:** in order of arrival to the queue; packets that arrive to a full buffer are either discarded, or a discard policy is used to determine which packet to discard among the arrival and those already queued



Scheduling Policies

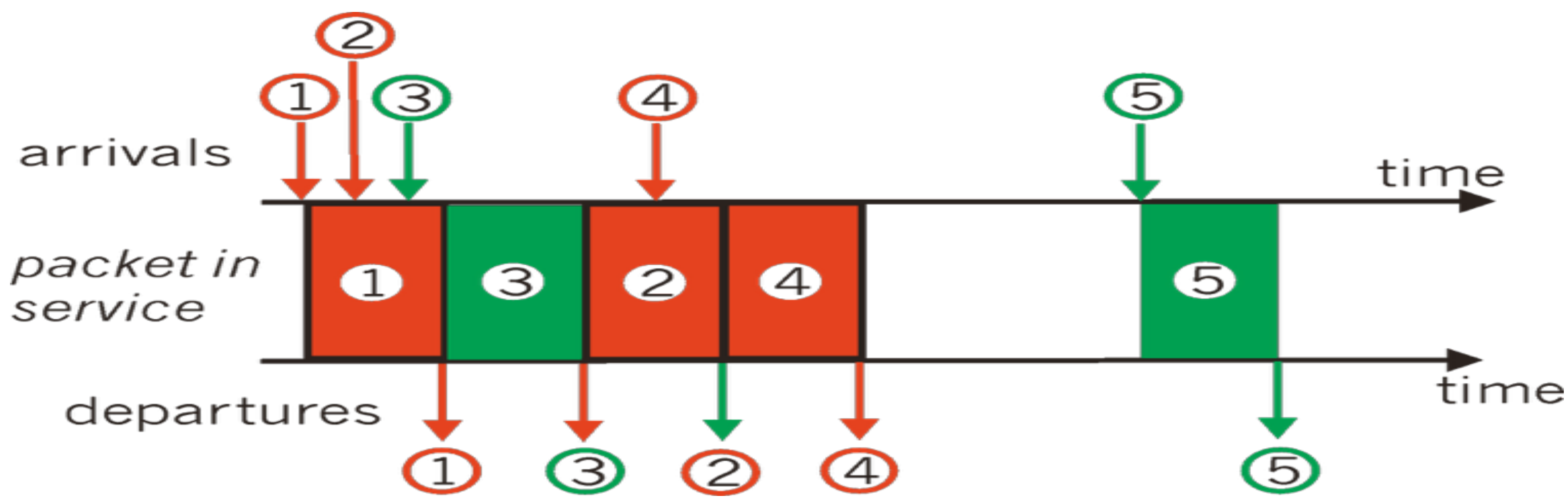
- Priority Queuing:
  classes have different priorities; class may depend on explicit marking or other header info, e.g. IP source or destination, TCP Port numbers, etc.
- Transmit a packet from the highest priority class with a non-empty queue
- Preemptive and non-preemptive versions





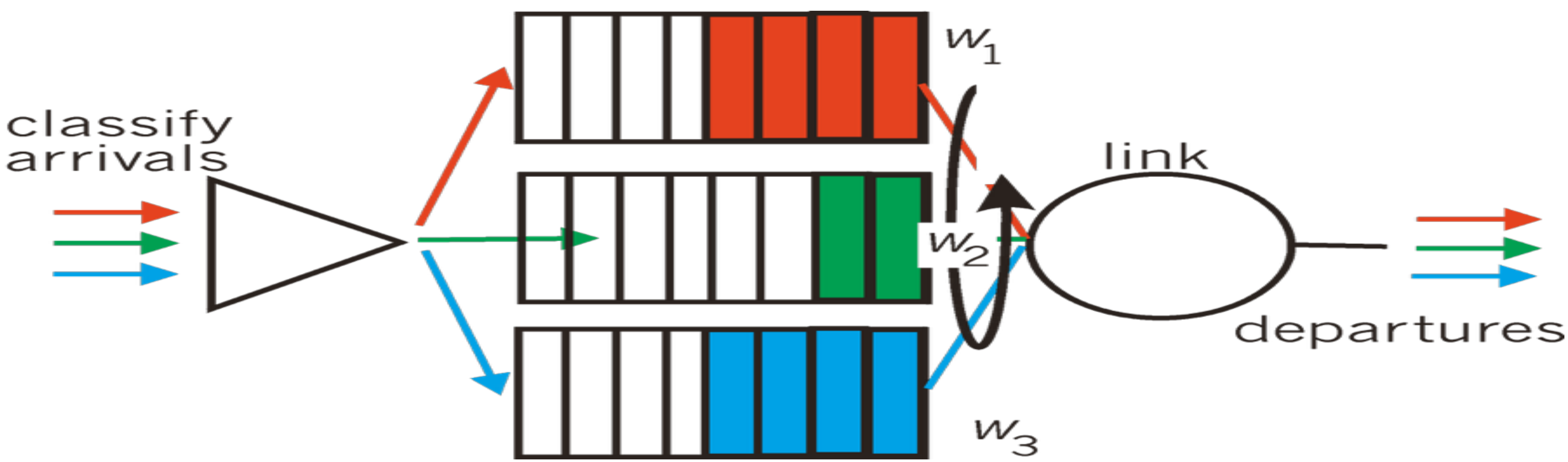
Scheduling Policies (more)

- Round Robin: scan class queues serving one from each class that has a non-empty queue



Scheduling Policies (more)

- Weighted Fair Queuing: is a generalized Round Robin in which an attempt is made to provide a class with a differentiated amount of service over a given period of time

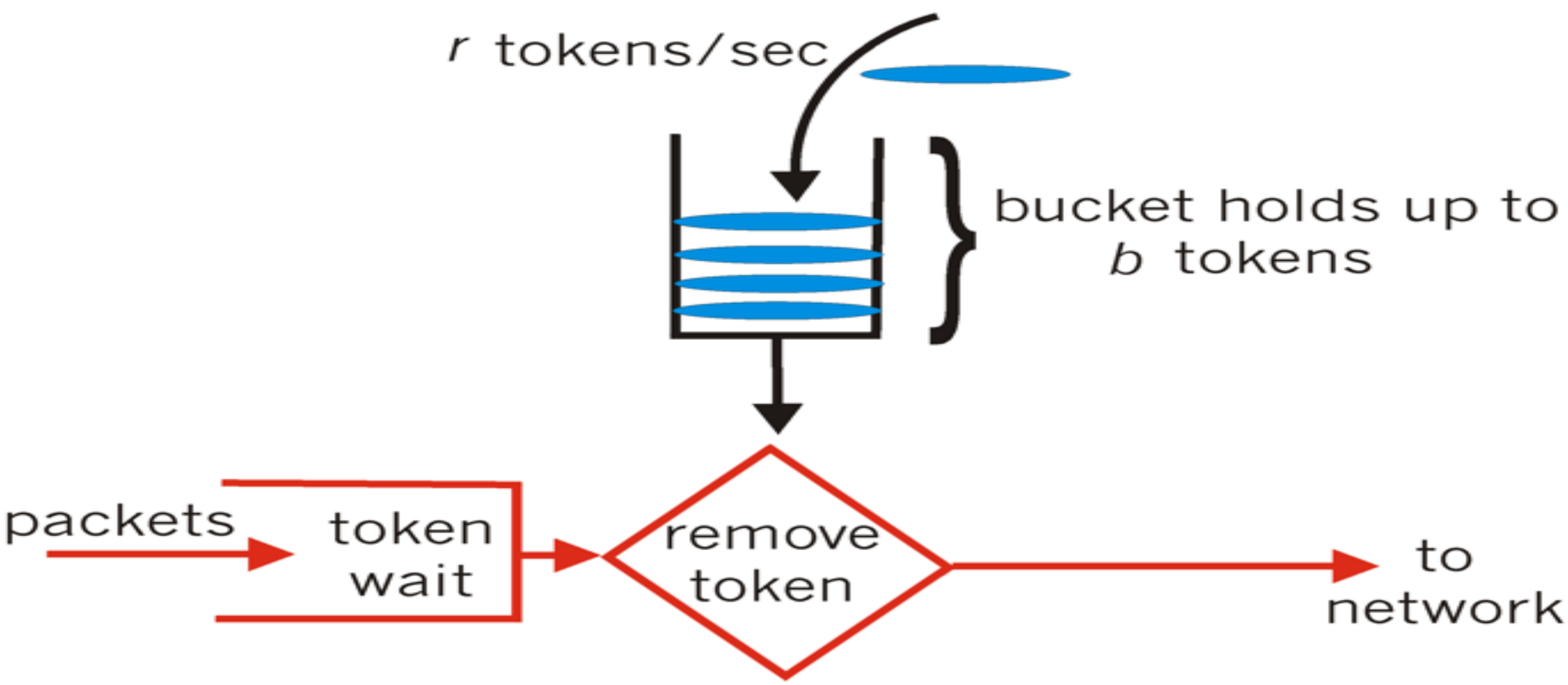


Policing Mechanisms

- Three criteria:
  - (Long term) **Average Rate** (100 packets/sec or 6000 packets/min), crucial aspect is the interval length
  - **Peak Rate**: e.g., 6000 p/min **Avg** and 1500 p/sec **Peak**
  - (Max.) **Burst Size**: Max. number of packets sent consecutively, i.e. over a short period of time

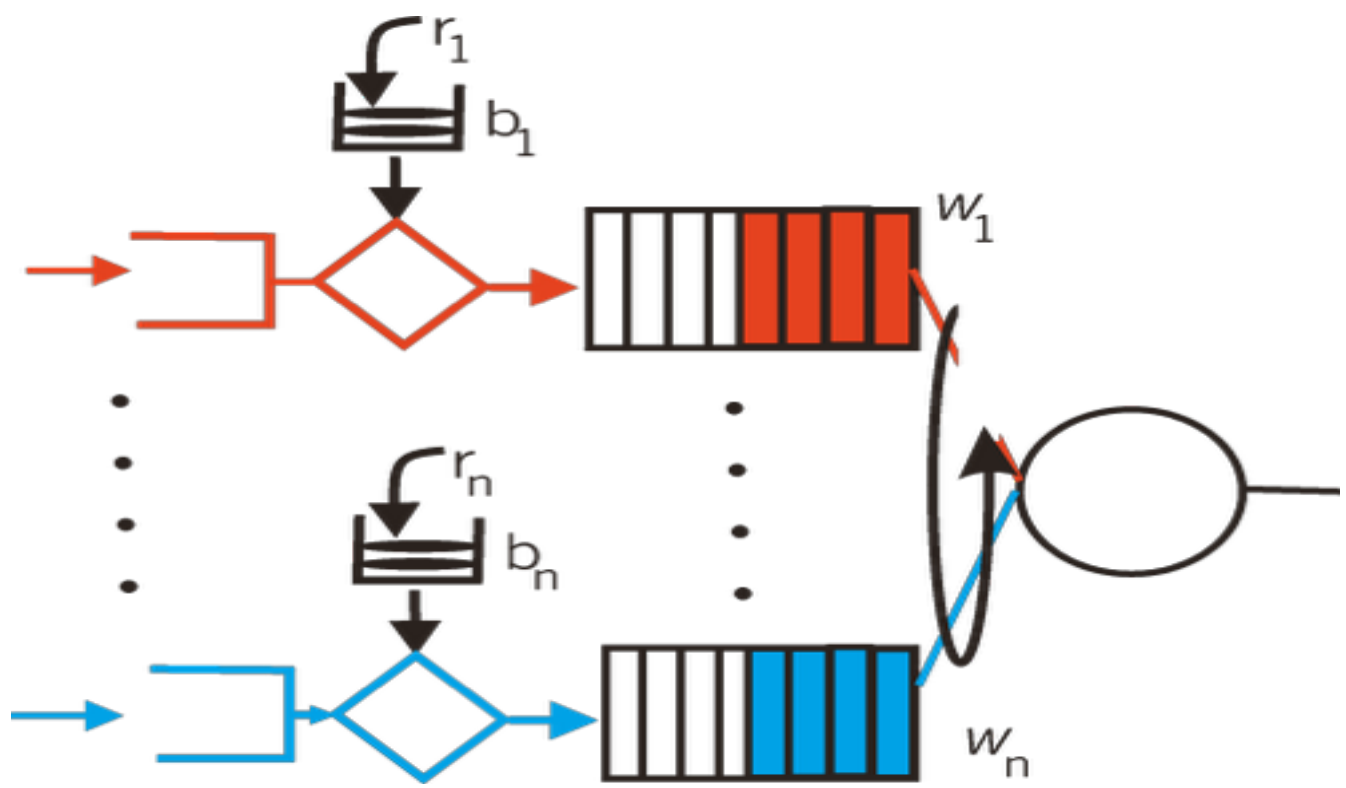
Policing Mechanisms

- **Token Bucket** mechanism, provides a means for limiting input to specified Burst Size and Average Rate.



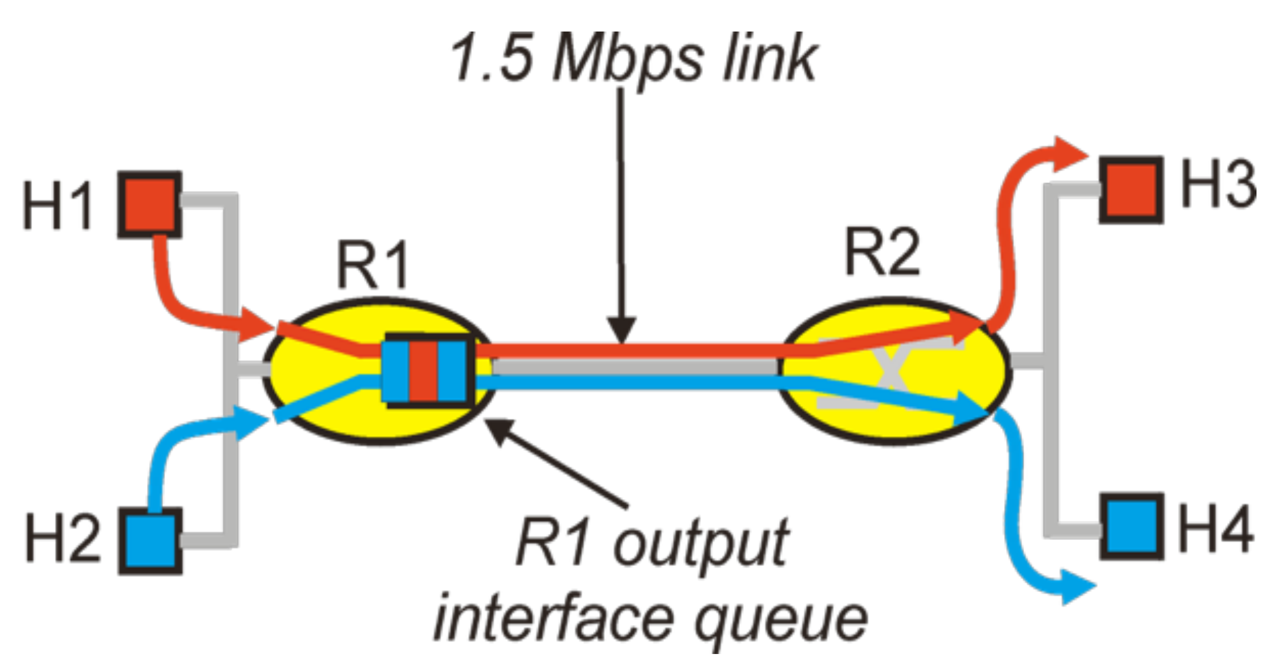
Policing Mechanisms (more)

- Bucket can hold *b* tokens; token are generated at a rate of *r token/sec* unless bucket is full of tokens.
- Over an interval of length *t*, the number of packets that are admitted is less than or equal to $(r\, t + b)$.
- Token bucket and WFQ can be combined to provide upper bound on delay.



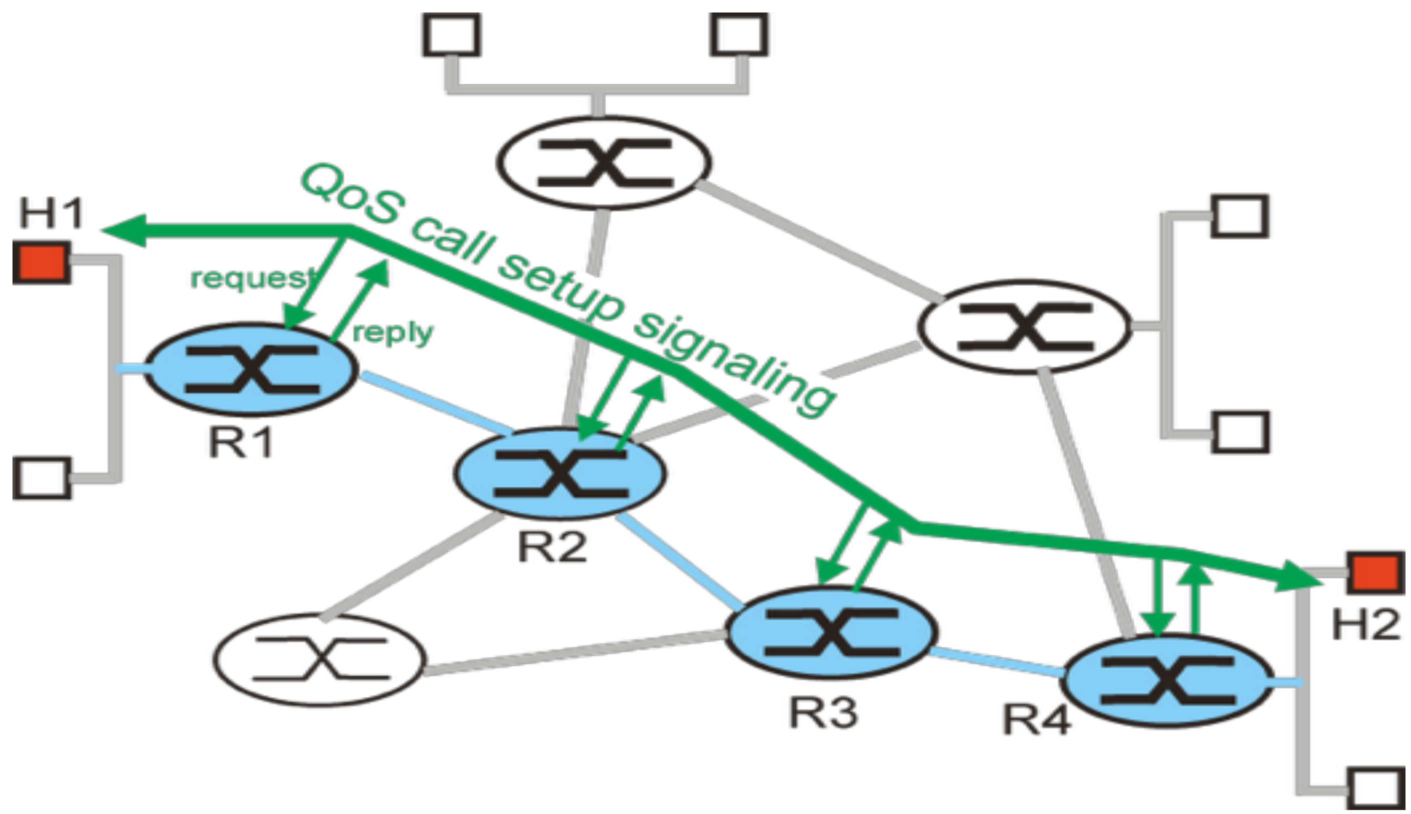
IETF IP QoS Efforts

- Policy based IP QoS Solutions
  - Integrated Services (RSVP protocol):
    - flow based
  - Differentiated Services (DiffServ byte settings):
    - packet based
  - Multi-Protocol Label Switching (MPLS):
    - flow+packet based
- IP Multicast and Anycast
- IPv6 QoS Support



Connection Oriented QoS

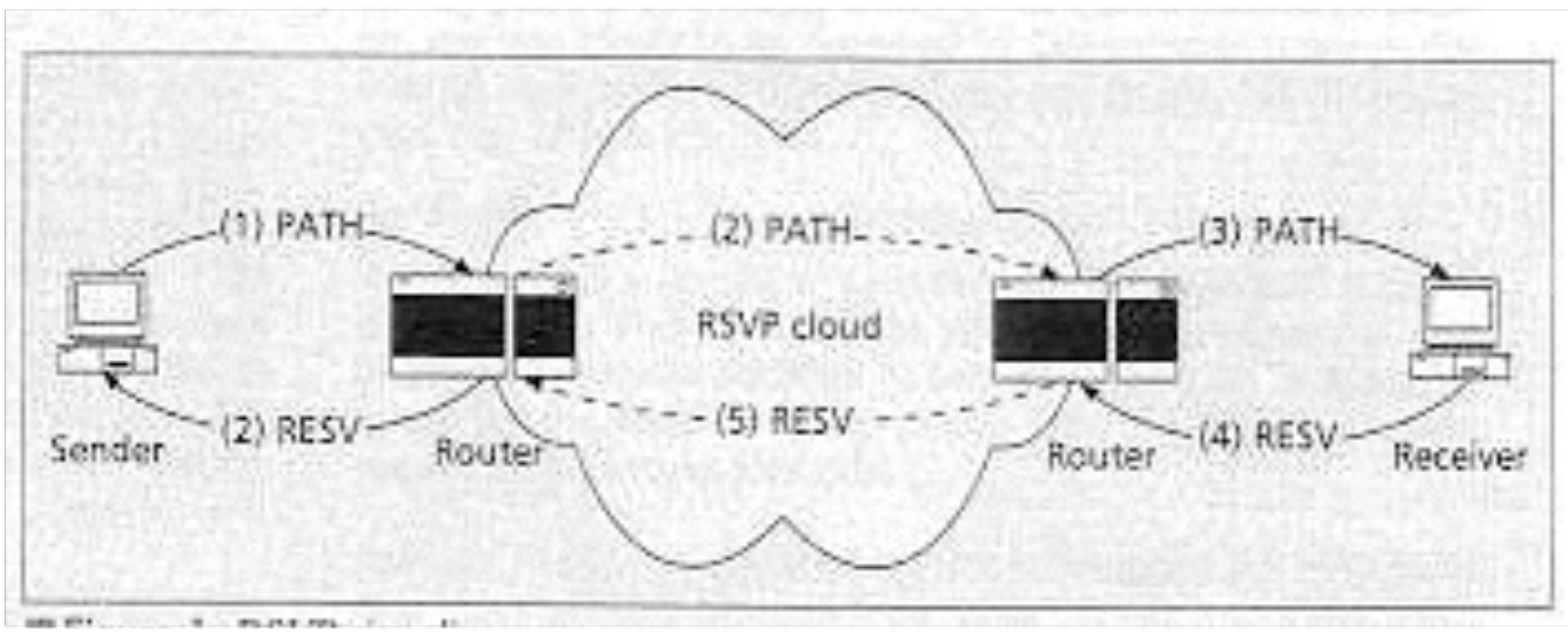
- **Integrated Services**: IETF RFC 1633
  - –Defined by RSVP requires resource reservation at each hop end-to-end for each IP packet flow, and end-to-end signaling along nodes in the path
  - –Reserve resources at the routers so as to provide QoS for specific user packet stream
  - –This architecture does not scale well (large amount of states)
  - –Many Internet flows are short lived, not worth setting up VC



Integrated Services / RSVP (警车开道)

- Sender sends a "**PATH**" message to the receiver specifying characteristics of traffic
  - every intermediate router along the path forwards the "PATH" message to the next hop determined by the routing protocol
- Receiver responds with "RESV" message after receiving "PATH". "RESV" requests resources for flow

RSVP
=
Resource reservation protocol



Connectionless QoS: IP Diff Serv

- Mark IP packet to specify treatment: IETF RFC 2474,
  - e.g., first class, business class, coach, standby

通行证机制

- Per Hop Behaviors (PHBs) based on network-wide traffic classes
- Flows are classified at the edge router based on rules, and are aggregated into traffic classes, allowing scalability
- Diff Serv uses the IP header TOS byte (first 6 bits), which is renamed the DS field
- Diff Serv defines code points (DSCP) for the DS field, DE (default) = 000000 = best effort, and EF (Expedited Forwarding) = 101110 = low latency, etc.

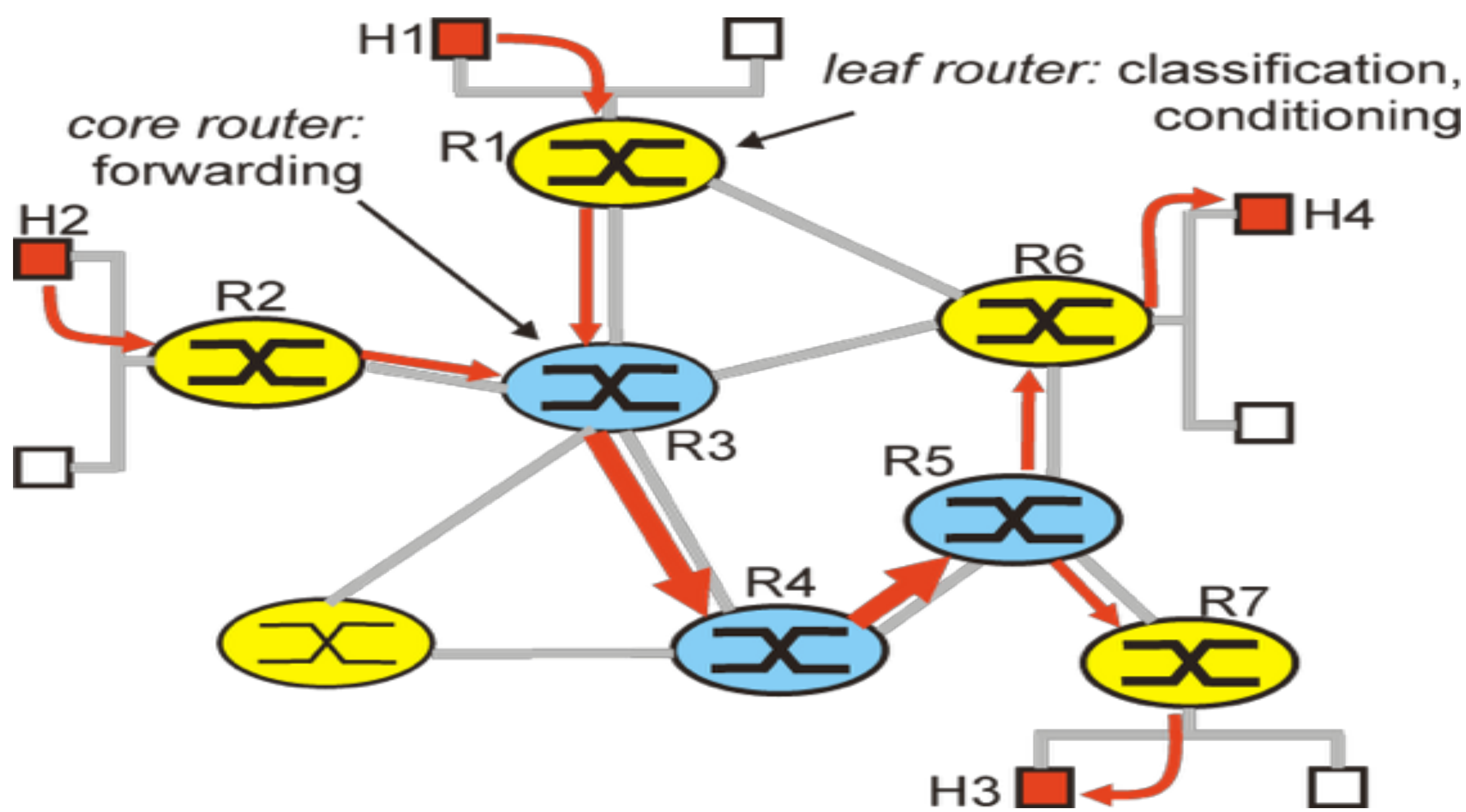
Differentiated Services

- Approach:
  - Only simple functions in the core, and relatively complex functions at edge routers (or hosts)
  - Do not define service classes, instead provides functional components with which service classes can be built

Edge Functions

- At DS-capable host or first DS-capable router

- **Classification**: edge node marks packets according to classification rules to be specified (manually by admin, or by some TBD protocol)
- **Traffic Conditioning**: edge node may delay and then forward or may discard



Core Functions

- **Forwarding**: according to "Per-Hop-Behavior" or PHB specified for the particular packet class; such PHB is strictly based on class marking (no other header fields can be used to influence PHB)

- BIG ADVANTAGE:

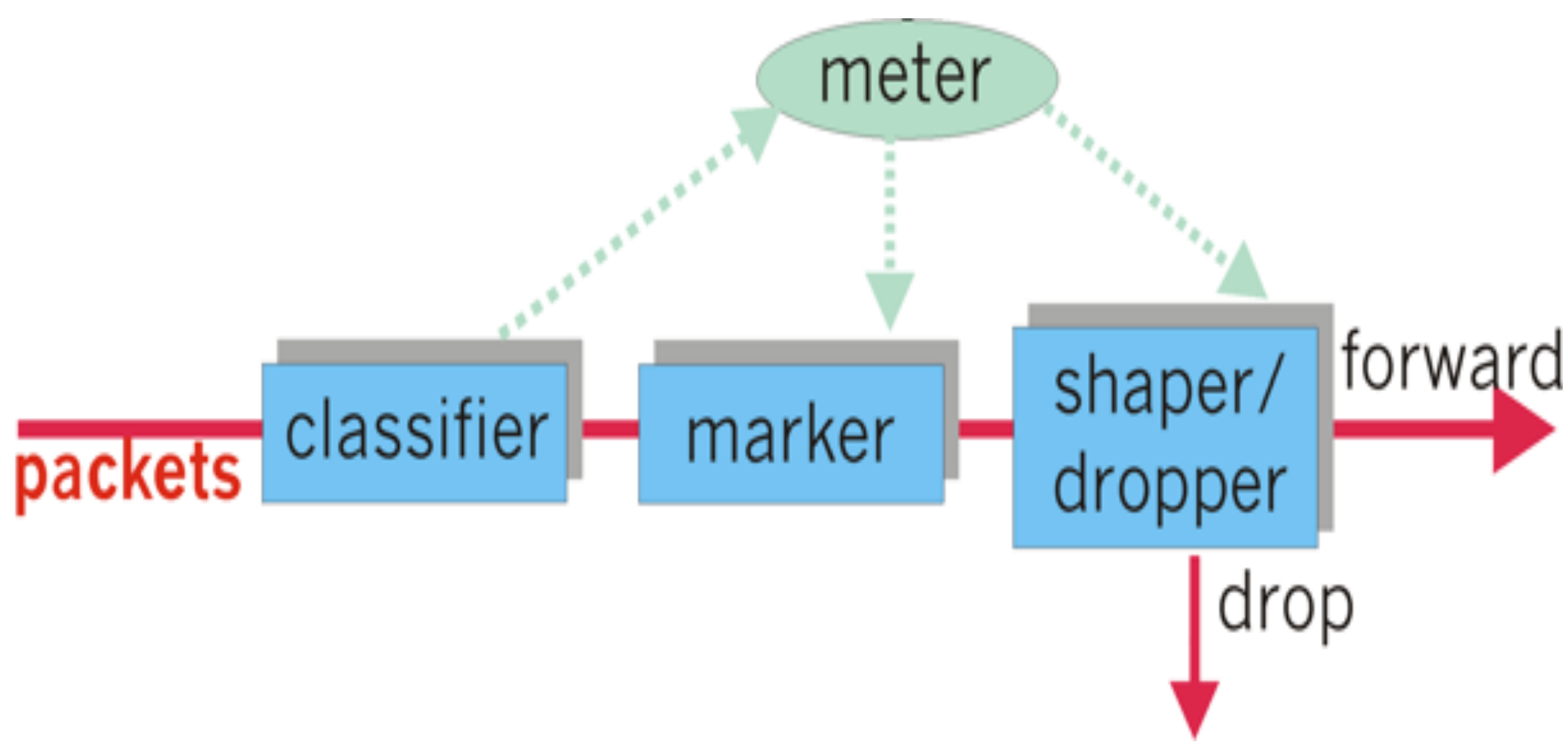
  No state info to be maintained by routers!

Classification and Conditioning

- Packet is marked in the Type of Service (TOS) in IPv4, and Traffic Class in IPv6
- 6 bits used for Differentiated Service Code Point (DSCP) and determine PHB that the packet will receive
- 2 bits are currently unused

Classification and Conditioning

- It may be desirable to limit traffic injection rate of some class; user declares traffic profile (e.g., rate and burst size);
- traffic is metered and shaped if non-conforming



QoS Priority Levels

Priority Level	Traffic Type
0	Best Effort
1	Background
2	Standard (Spare)
3	Excellent Load (Business Critical)
4	Controlled Load (Streaming Multimedia)
5	Voice and Video (Interactive Media and Voice) [Less than 100ms latency and jitter]
6	Layer 3 Network Control Reserved Traffic [Less than 10ms latency and jitter]
7	Layer 2 Network Control Reserved Traffic [Lowest latency and jitter]

IPv6 Support of QoS

- IPv6 Flow Labels provide support for Data Flows
  - Packet Prioritizing
    - sure that high priority traffic is not interrupted by less critical data

- IPv6 supports Multicast & Anycast
  - Multicast delivers data simultaneously to all hosts that sign up to receive it
  - Anycast allows one host initiate the efficient updating of routing tables for a group of hosts.

How should the Internet evolve to better support multimedia?

<u>Integrated services philosophy:</u>

- Change Internet protocols so that applications can reserve end-to-end bandwidth
  - Need to deploy protocol that reserves bandwidth
  - Must modify scheduling policies in routers to honor reservations
  - Application must provide the network with a description of its traffic, and must further abide to this description.
- Requires new, complex software in hosts & routers

<u>Differentiated services philosophy:</u>

- Fewer changes to Internet infrastructure, yet provide 1st and 2nd class service.
- Datagrams are marked.
- User pays more to send/receive 1st class packets.
- ISPs pay more to backbones to send/receive 1st class packets.

How should the Internet evolve to better support multimedia? (cont.)

Laissez-faire philosophy

- No reservations, no datagram marking
- As demand increases, provision more bandwidth
- Place stored content at edge of network:
  - ISPs & backbones add caches
  - Content providers put content in CDN nodes
  - P2P: choose nearby peer with content

Virtual private networks (VPNs)

- Reserve permanent blocks of bandwidth for enterprises.
- Routers distinguish VPN traffic using IP addresses
- Routers use special scheduling policies to provide reserved bandwidth.

Streaming Stored Audio & Video

Streaming stored media:

- Audio/video file is stored in a server
- Users request audio/video file on demand.
- Audio/video is rendered within, say, 10 s after request.
- Interactivity (pause, re-positioning, etc.) is allowed.

Media player:

- removes jitter
- decompresses
- error correction
- graphical user interface with controls for interactivity

- Plug-ins may be used to imbed the media player into the browser window