CS 414 – Multimedia Systems Design

Lecture 16 – Introduction to Multimedia Resource Management and Quality of Service

Klara Nahrstedt

Spring 2014

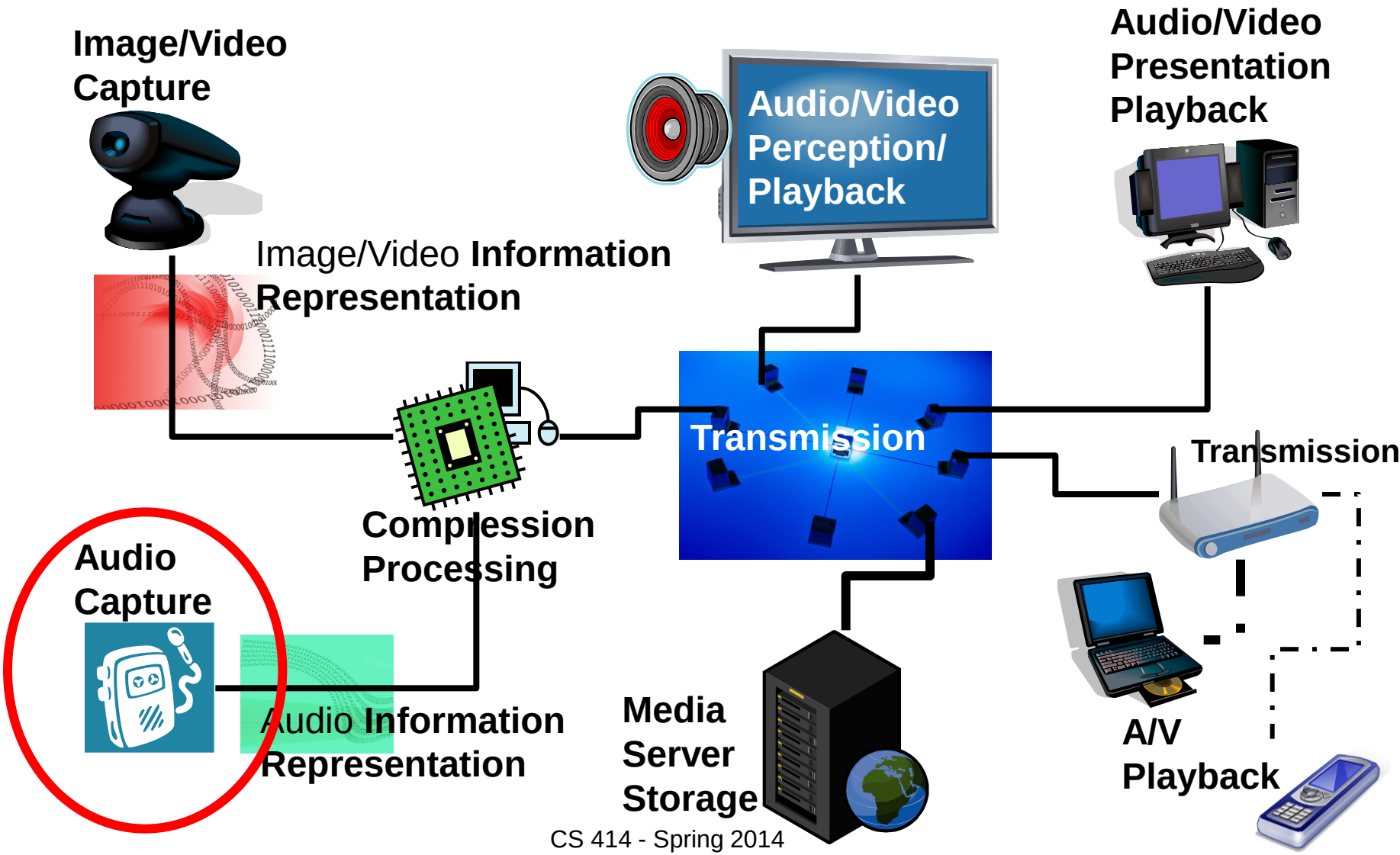
Administrative

- Reading: “Multimedia Systems”, Steinmetz and Nahrstedt, Springer 2004, Chapter 2
- Reading: “Multimedia Systems: Algorithms, Standards, and Industry Practices”, Havaldar and Medioni, Chapter 11
- HW1 posted on Monday, February 24.
  - HW1 due on Monday, March 3.

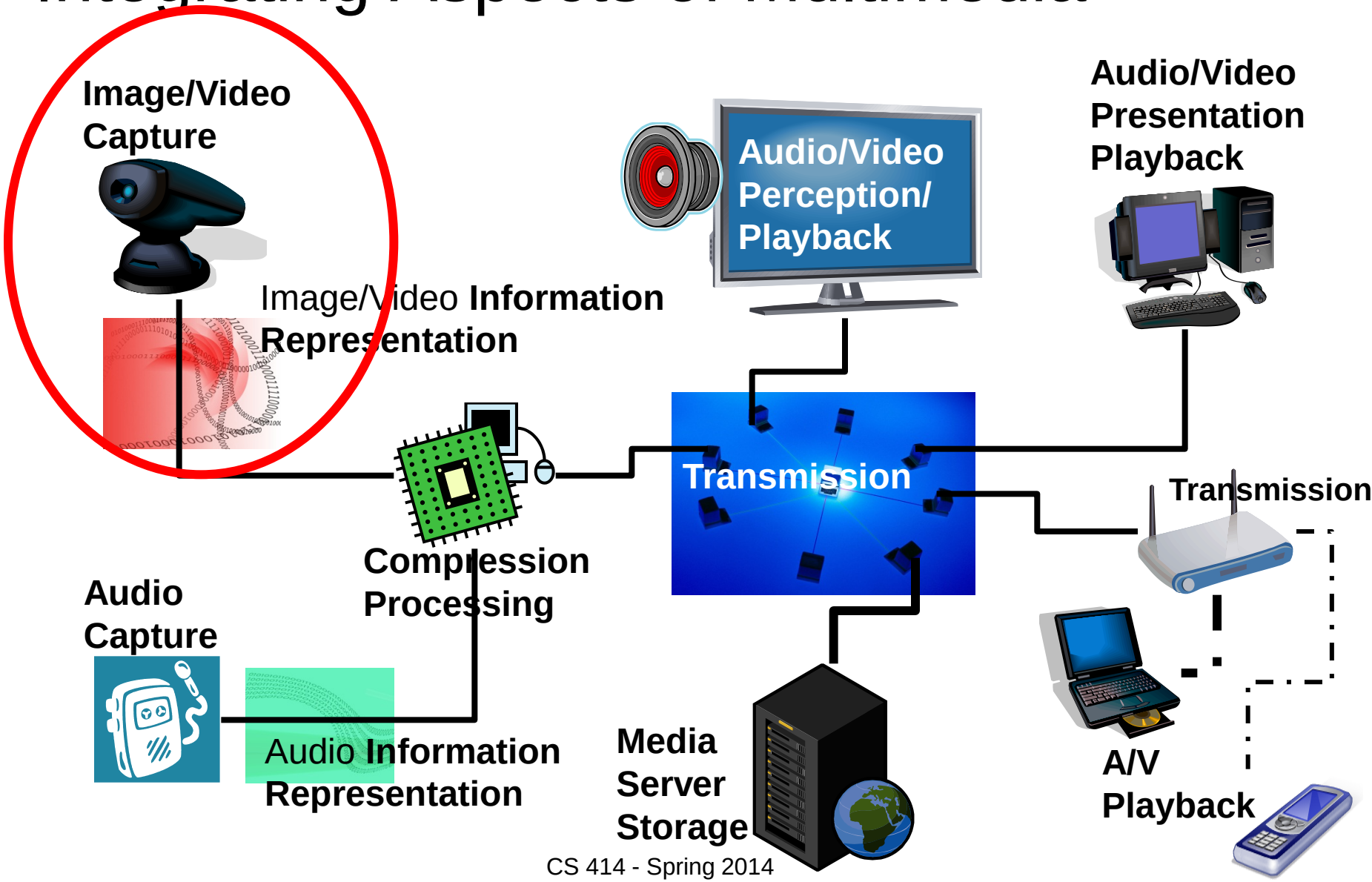
Outline

- AV Requirements - Real-time
- AV Requirements on Multimedia Networks and Operating Systems
- Resource management
  - Resources
  - Quality of Service (QoS) Concept
  - Operations

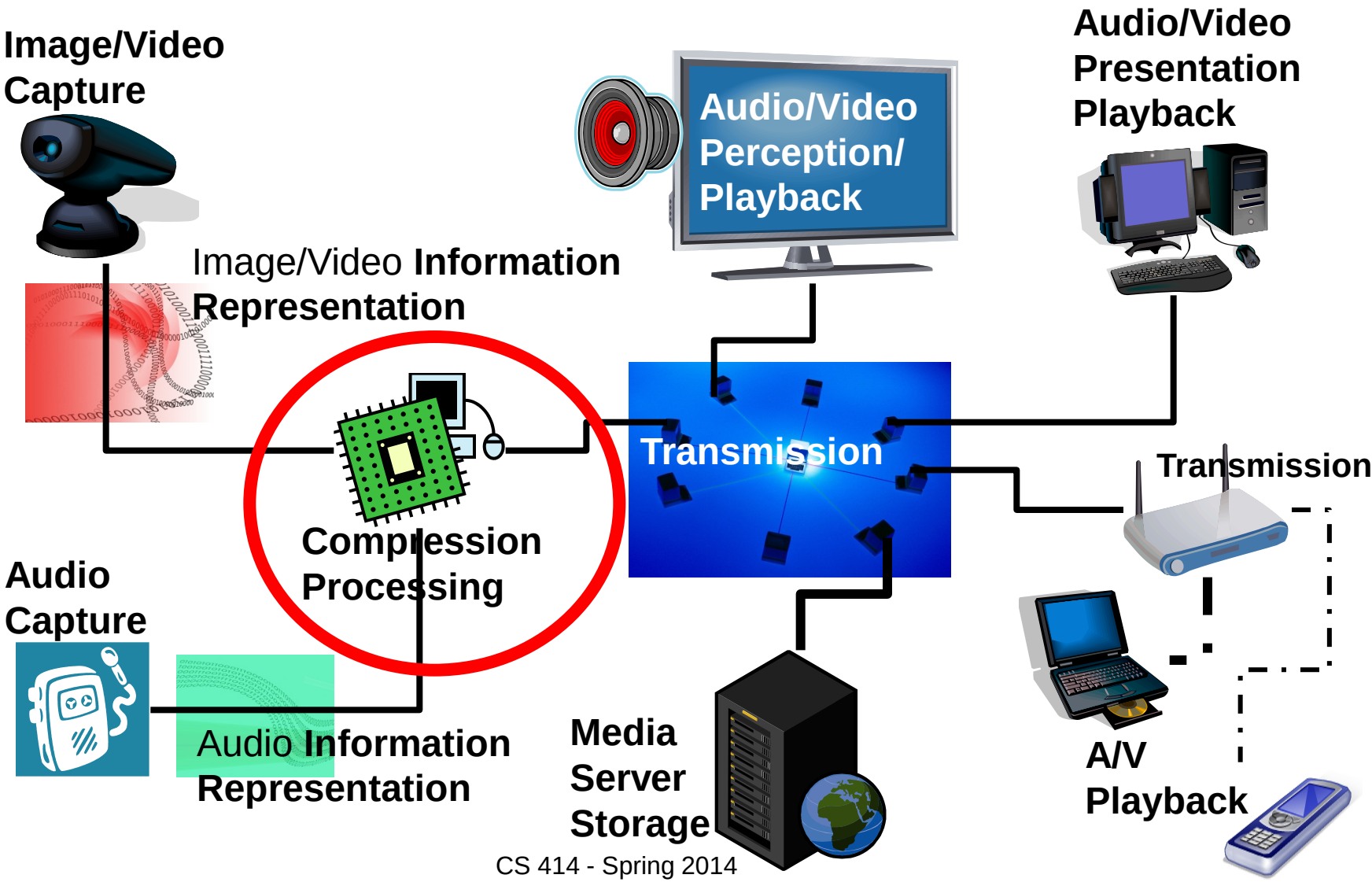
Integrating Aspects of Multimedia



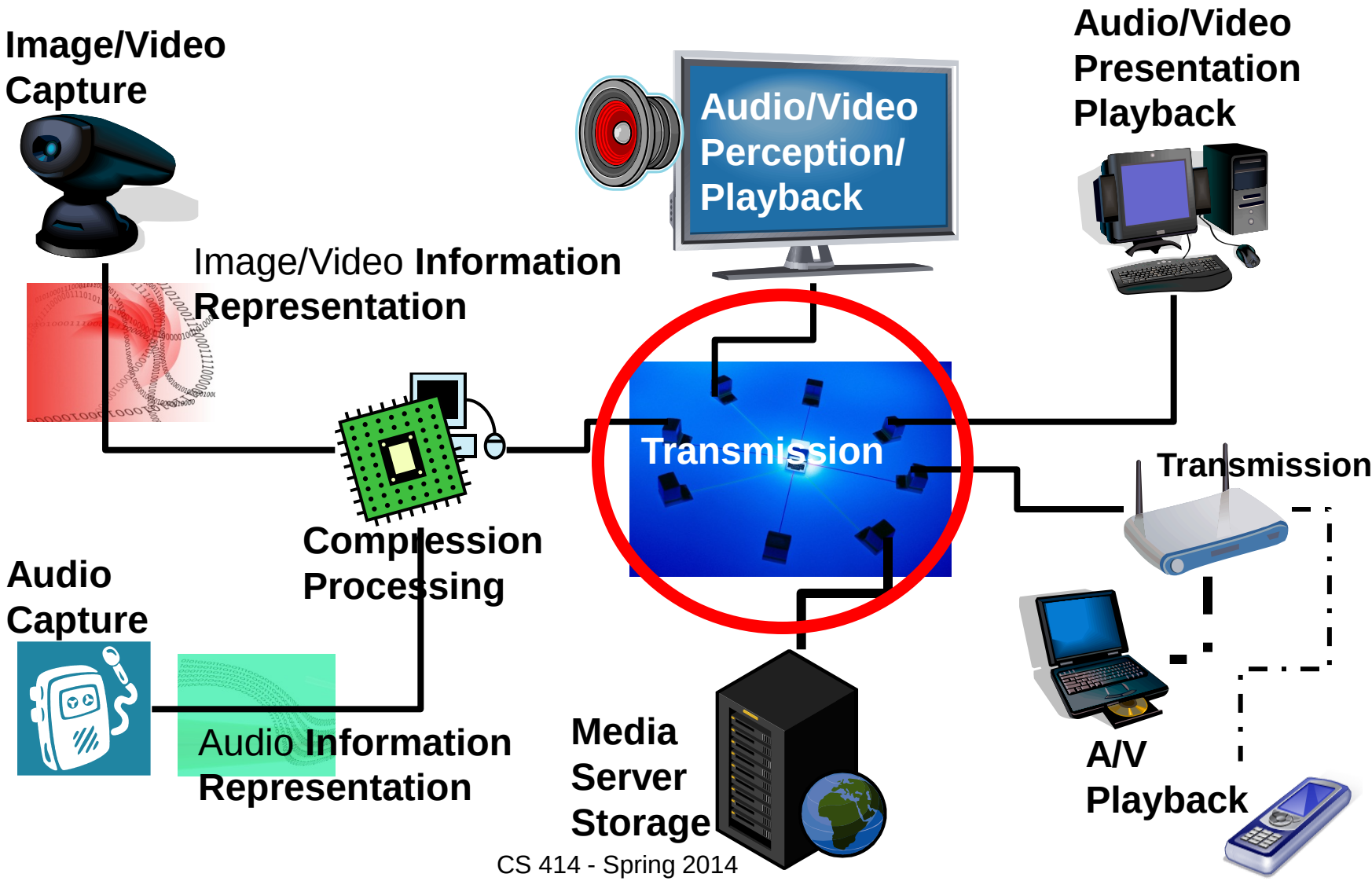
Integrating Aspects of Multimedia



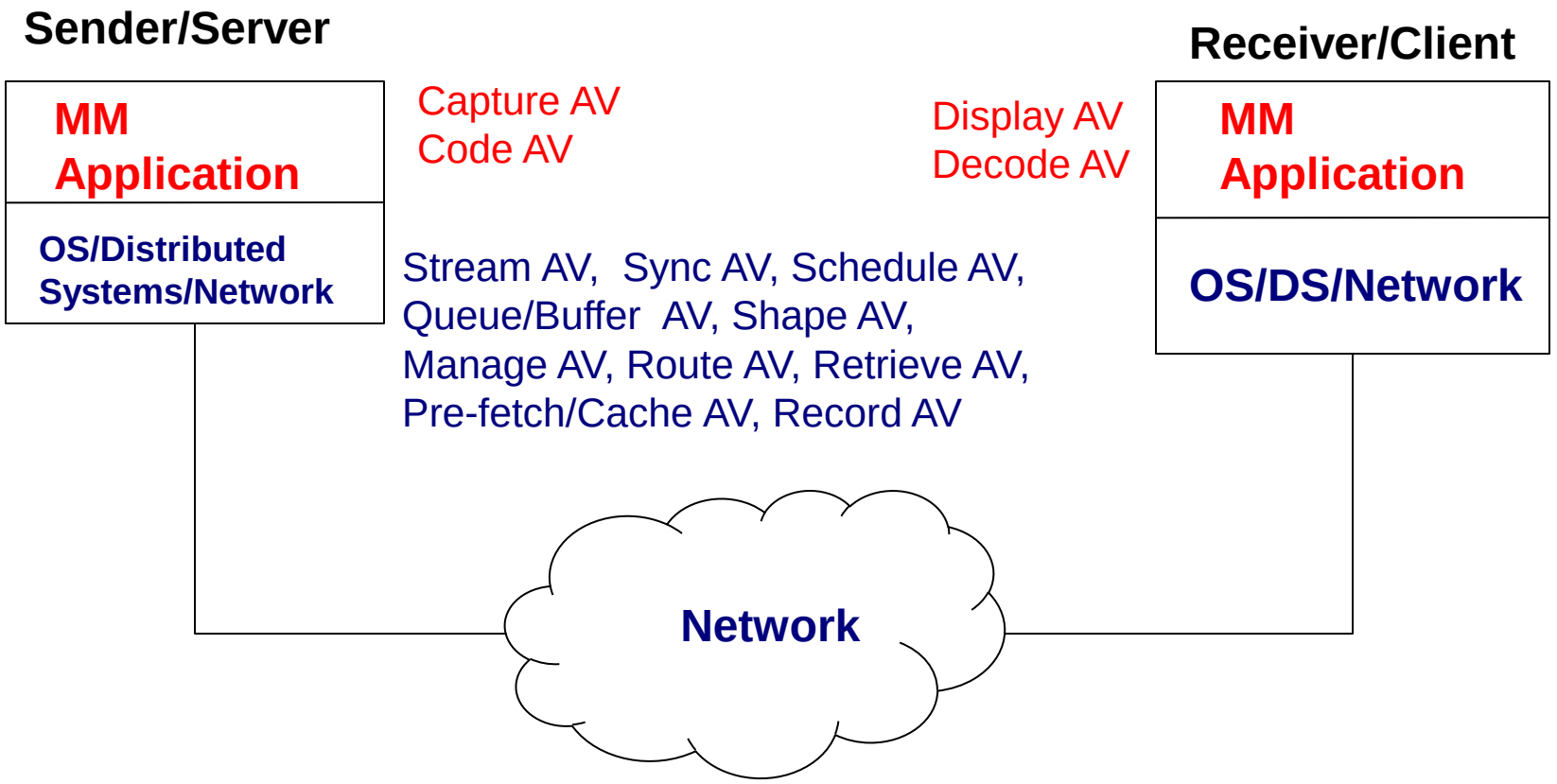
Integrating Aspects of Multimedia



For Next Five Weeks we will cover Transmission/Networks of Multimedia



Multimedia Distributed System and Network

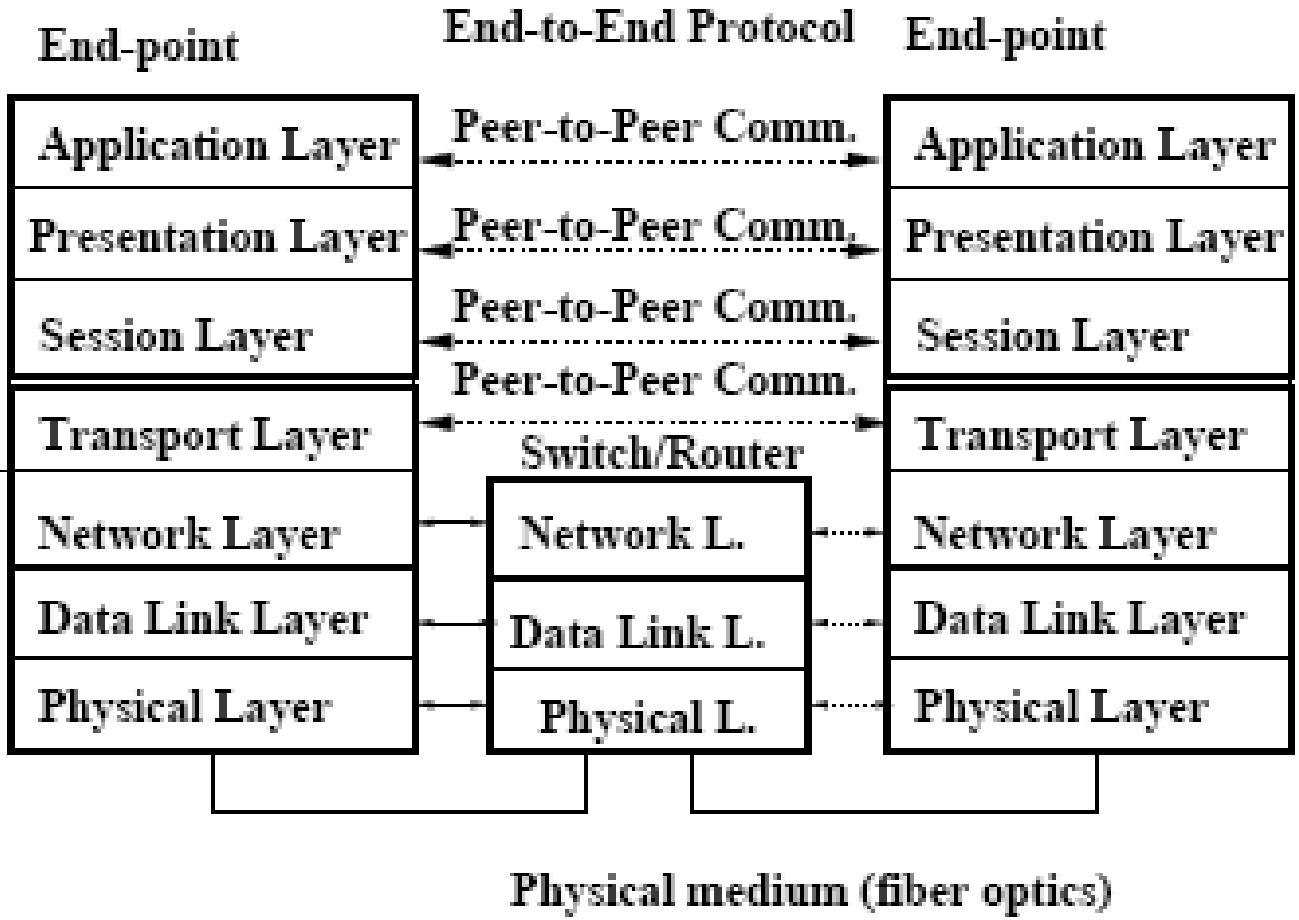


Network Model = OSI (Open System Interconnection) Layering Standard

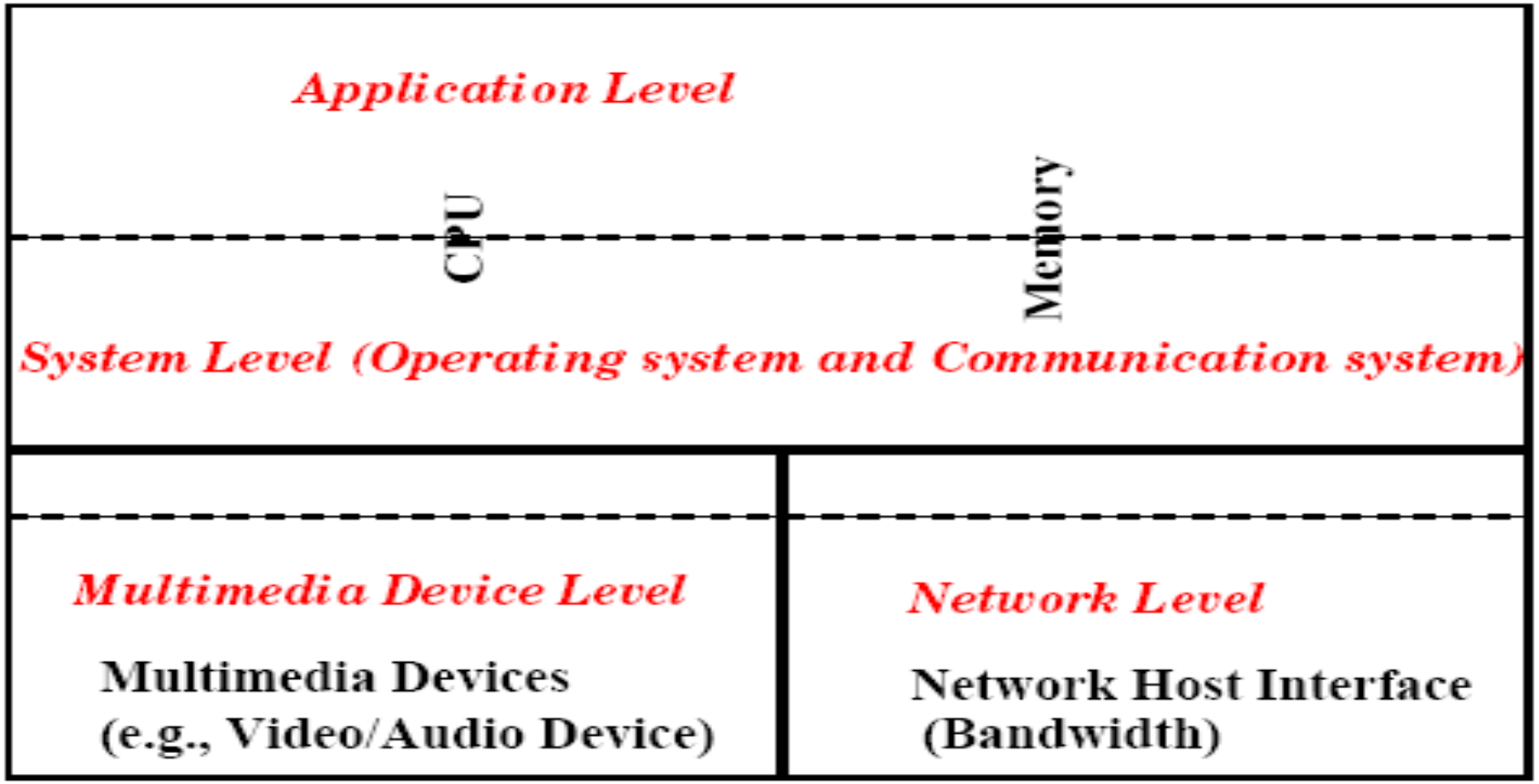
VOD Services
(Video Retrieval And Video Playback)

Peer-to-Peer Streaming

Network QoS/Resource Management



Layered Partition of Multimedia Systems with respect to Required Resources and Individual Services



AV Requirements: Real-Time and Deadlines

- Real-time system – system in which correctness of computation depends not only on obtaining the right results, but also providing them on time
  - Examples: control of temperature in a chemical plant; control of a flight simulator
- Deadline – represents the latest acceptable time for the result delivery
  - Soft deadlines versus hard deadlines

AV Requirements: Real-Time and Multimedia

- Difference between RT requirements for traditional RT systems and Multimedia systems
  - Soft deadlines versus hard deadlines
  - Periodic behavior versus random behavior
  - Bandwidth requirements

AV Requirements on MM Systems and Networks

- Transport system – guaranteed delivery with respect to metrics such as delay, reliability, bandwidth requirements
- OS process management – real-time processing of continuous data, communication and synchronization between processes/ threads

AV Requirements on MM Systems and Networks (2)

- Memory/Buffer management – guaranteed timing delay and efficient data manipulation
- File system/Media Servers – transparent and guaranteed continuous retrieval of audio/video
- Device management – integration of audio and video

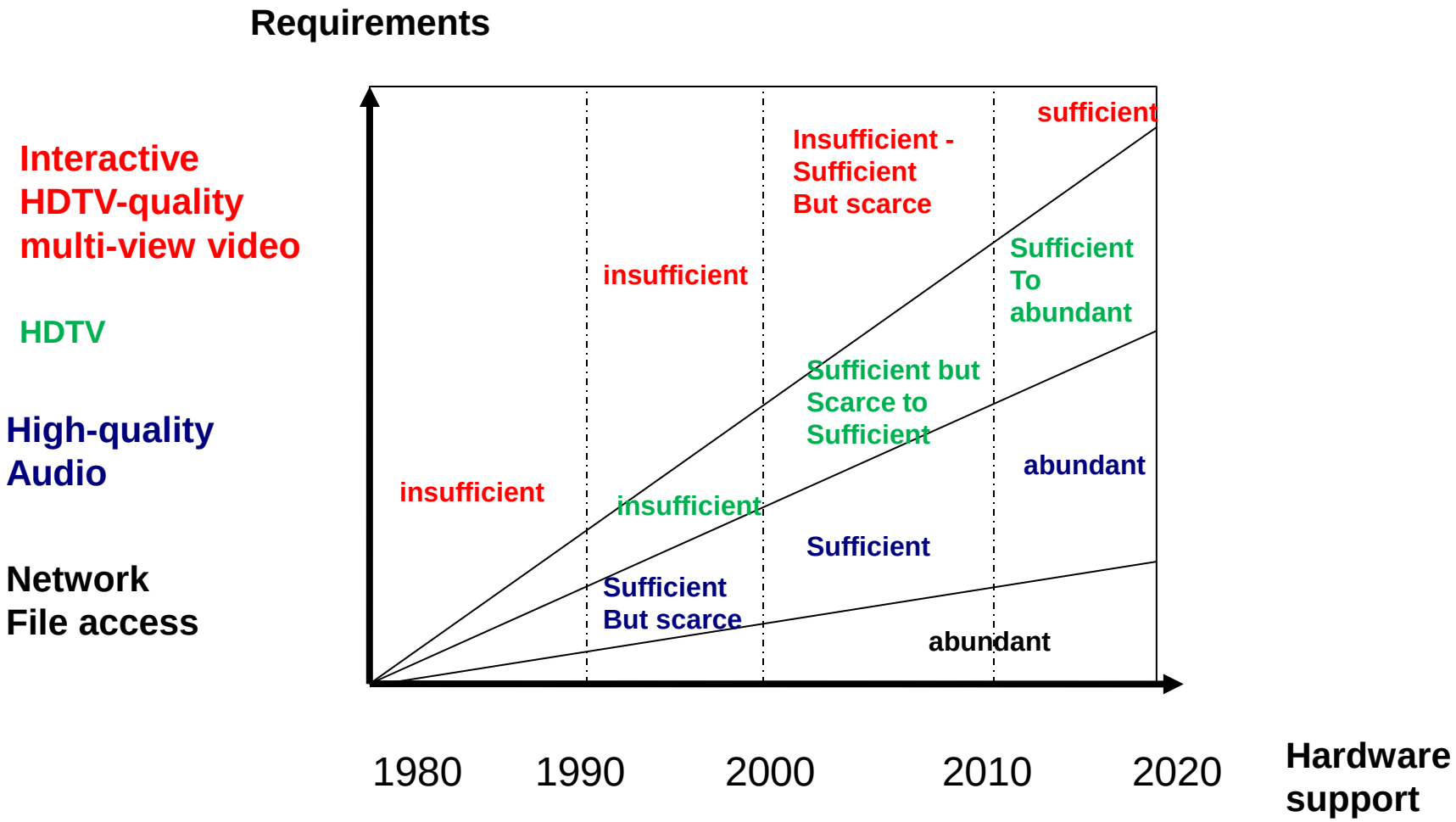
Result of AV Requirements

- **Need Resource Management** to coordinate
  - Transport/Network Resources,
  - CPU/OS Resources
  - Memory/Buffer Resources
  - Storage/Disk Resources
  - Device Resources

Resource Management (Why do we need resource management?)

- Limited capacity in digital distributed systems despite data compression and usage of new technologies
- Need adherence for processing of continuous data by every hardware and software component along the data path
- Competition for resources exist in an integrated multimedia system

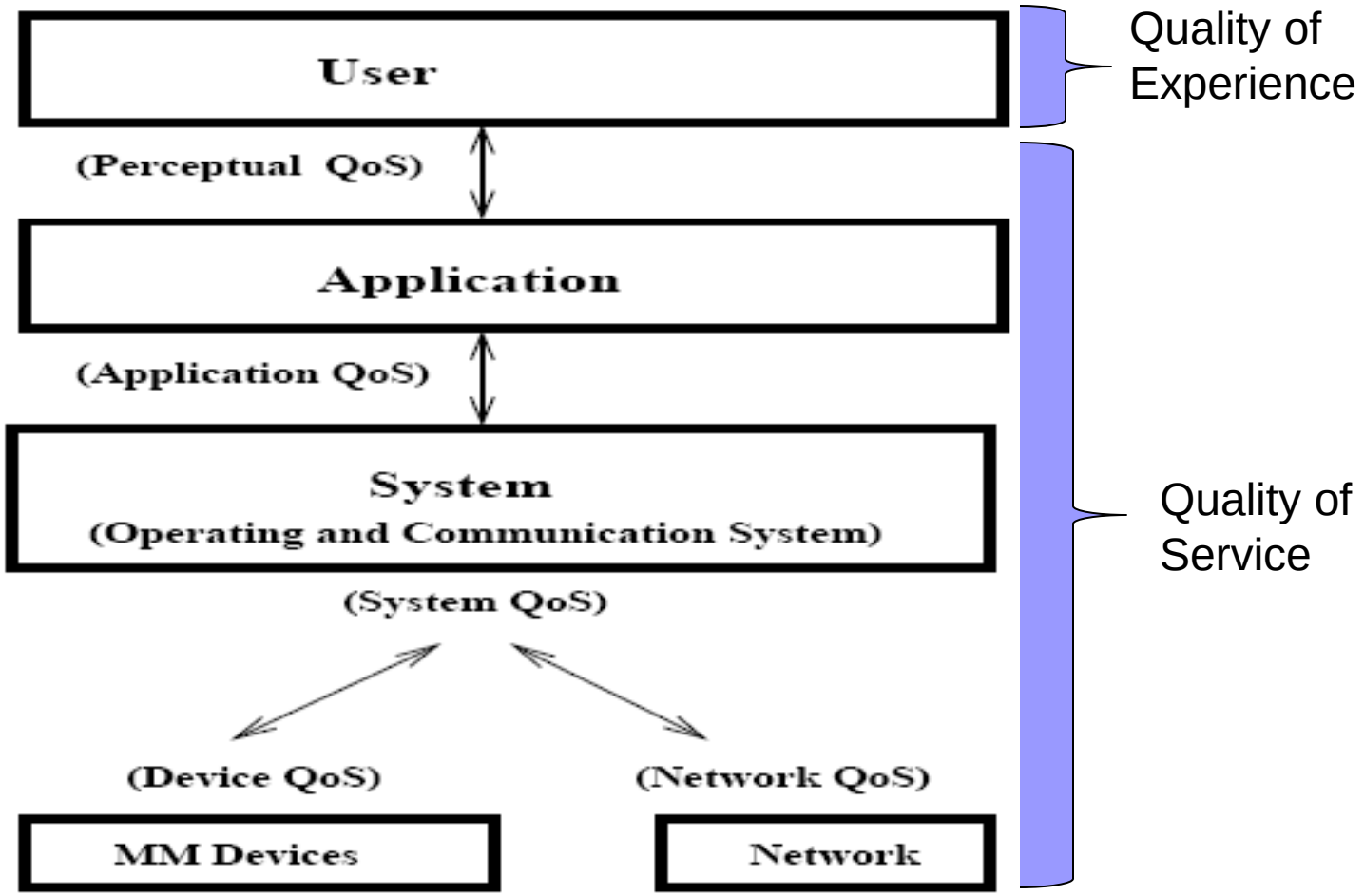
Window of Resources



Quality of Service (How to parameterize services?)

- To manage resources, we need services over resources
  - to schedule AV data, to shape access for AV data, to process AV data, to move AV data, etc.
- Multimedia systems consist of set of AV-specific services
  - **Processing (media-related) services**: retrieve audio/video, record video/audio, compress audio/video, fast forward video, rewind video
  - **Transport (network) services**: Stream video, fast forward video, rewind video
- To provide multimedia services, services get parameterized with quality levels called Quality of Service
- **QoS parameters versus performance metrics!!**

Layered Model for QoS



Application AV QoS Parameters

- **QoS for Audio service:**
  - **Sample rate** – 8000 samples/second (8KHz), 44.1 KHz
  - **Sample resolution** – 8 bits per sample, 16 bits per sample
- **QoS for Video service:**
  - **Video frame rate** – 25 frames per second, 30 frames per second
  - **Frame Period** – 40 ms, 30 ms, 25 ms, …
  - **Frame resolution** – 320x240 pixels, 640x480 pixels, 1920x1080 pixels, …
  - **Pixel resolution** – 24 bits per pixel, 8 bits per pixel
  - **Frame size** – 64KB
  - **Compression rate** – 8:1

Network QoS

- **Bandwidth** – Rate of data transfer, Bit Rate
  - e.g., 1 Gbps (Ethernet throughput) – level 1
  - e.g., 100 Mbps (WiFi throughput) – level 2
  - e.g., 128 kbps (ISDN throughput) – level 3
  - measured in bits per second
- **Throughput** – rate of successful message delivery over communication channel
  - Measured in packets per second, data packets per time slot, or bits per second
    - 30 packets per second; 128 kbps, 10 packets per time slot

Network QoS

- **Connection setup time**
  - time how long it take to connect the sender and receiver
  - e.g., 50 ms, 10 ms, …
- **Error Rate**
  - Measures the total number of bits (packets) that were corrupted or incorrectly received compared with the total number of transmitted bits (packets)
    - Bit Error Rate (BER) – at physical/MAC layer
      - In fiber optics, bit error rate (BER) is of the order of 10-8 to 10-12.
      - In satellite networks, BER is of the order 10-7
    - Packet Error Rate (PER) – at IP/transport/application layer – also called Packet Loss Rate

Network QoS

- **Delay**
  - Latency
    - End-to-end delay in telecommunication
  - Response time
    - Round-trip delay in telecommunication
- **End-to-End Delay**
  - time interval from the time packet is sent from the sender until the time it is received at the receiver ($T_{receive} - T_{send}$)
  - e.g., 80 ms, 100 ms, 160 ms

Network QoS

- **Response Time**
  - Measured as **round-trip delay** and is the total time required for sender to send a packet and receive an acknowledgement from the receiver. It can be described as sum of network delay and interface delay.
    - Network delay – composed of transit delay and transmission delay
      - Transit delay is caused by time needed to send data on a physical connection between sender and receiver
      - Transmission delay is time needed to transmit packet through network as result of processing delays (e.g., look up routing tables)
    - Interface delay – incurred between the time a sender is ready to begin sending and the time a network is ready to accept and transmit the data (due to traffic policing and shaping)

Other QoS Parameters

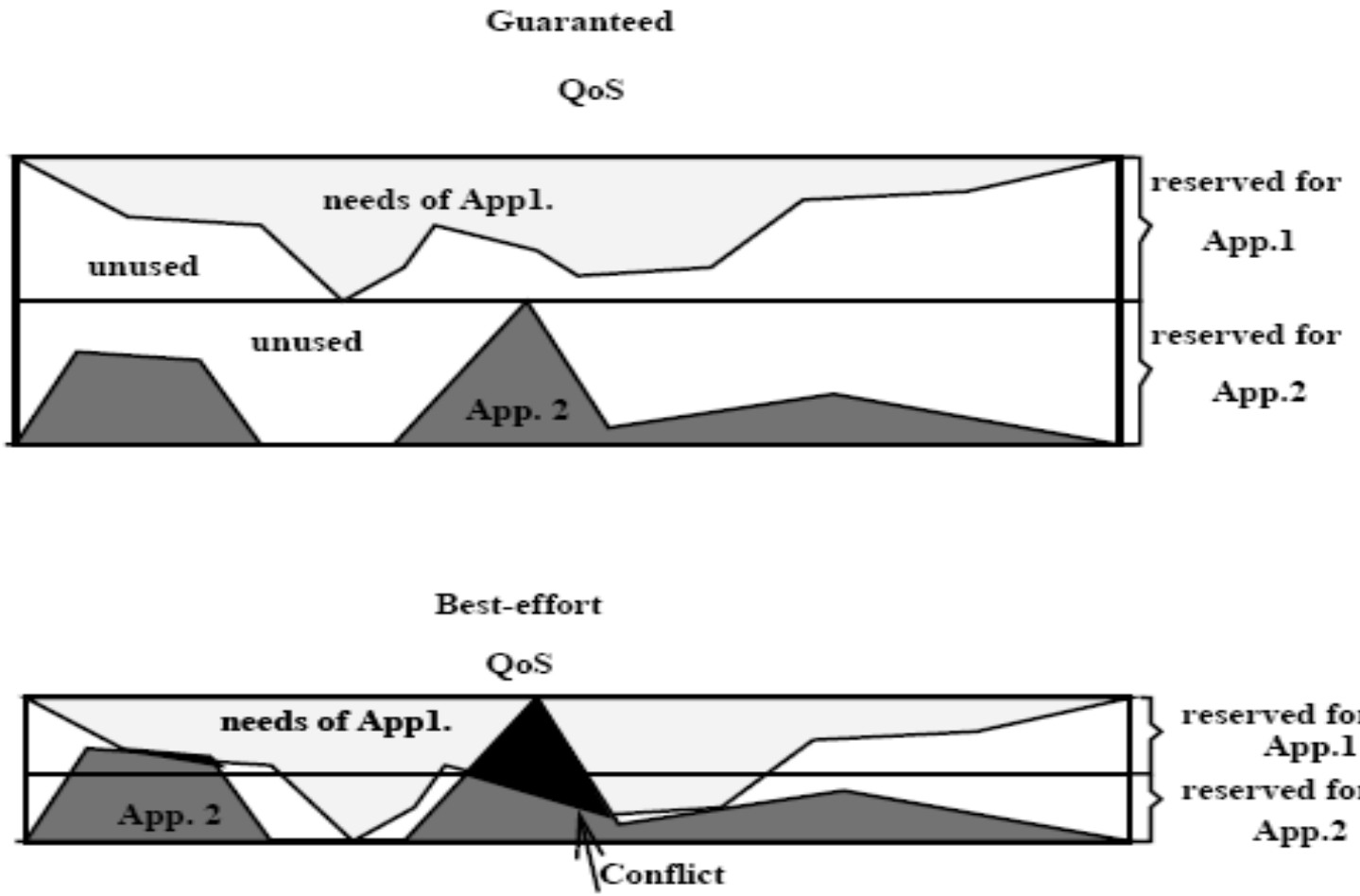
- Jitter
  - Undesired deviation from true periodicity in telecommunication
    - Also called packet delay variation – important QoS factor in assessment of network performance
  - **Packet jitter** – variation in latency as measured in the variability over time of the packet latency across network.

QoS Classes

- Guaranteed Service Class
  - QoS guarantees are provided based on deterministic and statistical QoS parameters
- Predictive Service Class
  - QoS parameter values are estimated and based on the past behavior of the service
- Best Effort Service Class
  - There are no guarantees or only partial guarantees are provided

QoS Classes (cont.)

QoS Class determines: (a) reliability of offered QoS, (b) utilization of resources

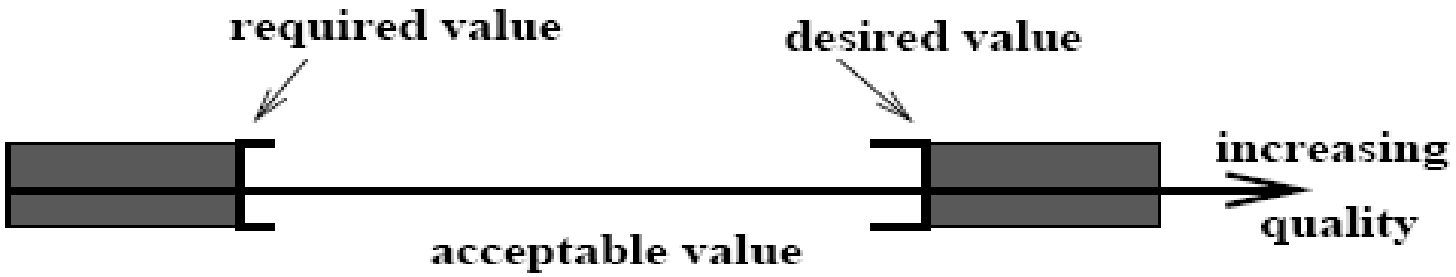


Deterministic QoS Parameters

- Single Value: QoS_1 – average *(QoS_{ave})*, contractual value, threshold value, target value
  - **Throughput – 10 Mbps**
- Pair Value: *<QoS_1, QoS_2>* with

  QoS_1 – required value; QoS_2 – desired value

  <QoS_{avg},QoS_{peak}>; <QoS_{min}, QoS_{max}>
  - **Throughput - <8,12> Mbps**



Deterministic QoS Parameter Values

- Triple of Values <QoS_1, QoS_2, QoS_3>
  - QoS_1 – best value
  - QoS_2 – average value
  - QoS_3 – worst value
- Example:
  - <QoS_{peak}, QoS_{avg}, QoS_{min}>, where QoS is network bandwidth
  - Throughput <12, 10, 8> Mbps

Guaranteed QoS

- We need to provide 100% guarantees for QoS values (hard guarantees) or very close to 100% (soft guarantees)
- Current QoS calculation and resource allocation are based on:
  1. Hard upper bounds for imposed workloads
  2. Worst case assumptions about system behavior

1. **Advantages:** QoS guarantees are satisfied even in the worst case case (high reliability in guarantees)
2. **Disadvantage:** Over-reservation of resources, hence needless rejection of requests

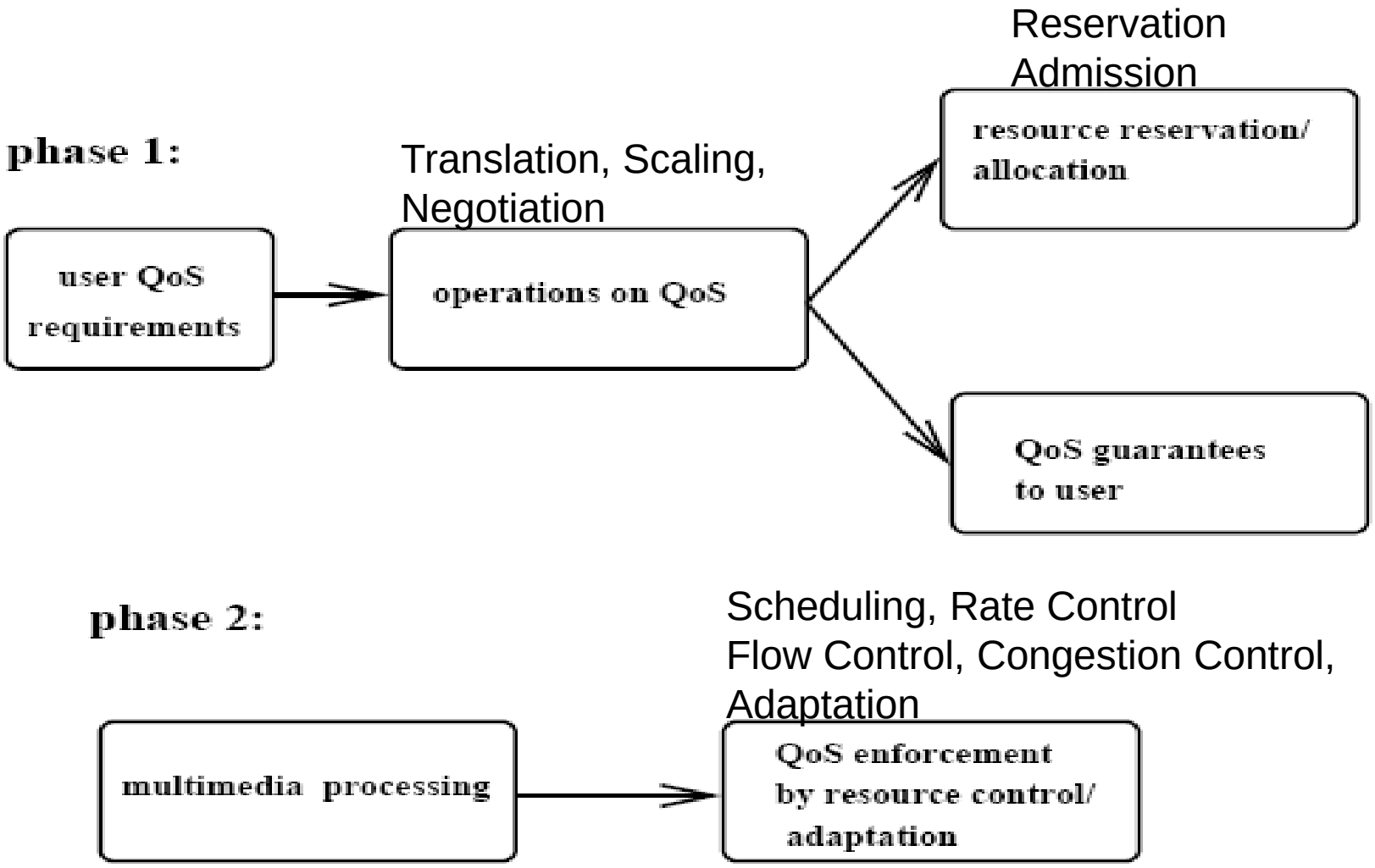
Predictive QoS Parameters

- We utilize QoS values $(QoS_1, ..QoS_i)$ and compute average
  - QoS_{bound} step at $K>i$ is $QoS_K = 1/i * \sum_j QoS_j$
- We utilize QoS values $(QoS_1, , QoS_i)$ and compute maximum value
  - $QoS_K = max_{j=1,...i} (QoS_j)$
- We utilize QoS values $(QoS_1, , QoS_i)$ and compute minimum value
  - $QoS_K = min_{j=1,...i} (QoS_j)$

Best Effort QoS

- No QoS bounds or possible very weak QoS bounds
- **Advantages**: resource capacities can be statistically multiplexed, hence more processing requests can be granted
- **Disadvantages**: QoS may be temporally violated

Relation between QoS and Resources



Conclusion

- QoS – an important concept in multimedia systems
- Very different types of QoS parameters and values
- Important relation between QoS and Resources
- Need to understand operations on QoS and their impact on resource management