CS 414 – Multimedia Systems Design

Lecture 13 – Quality of Service Concepts(Part 2)

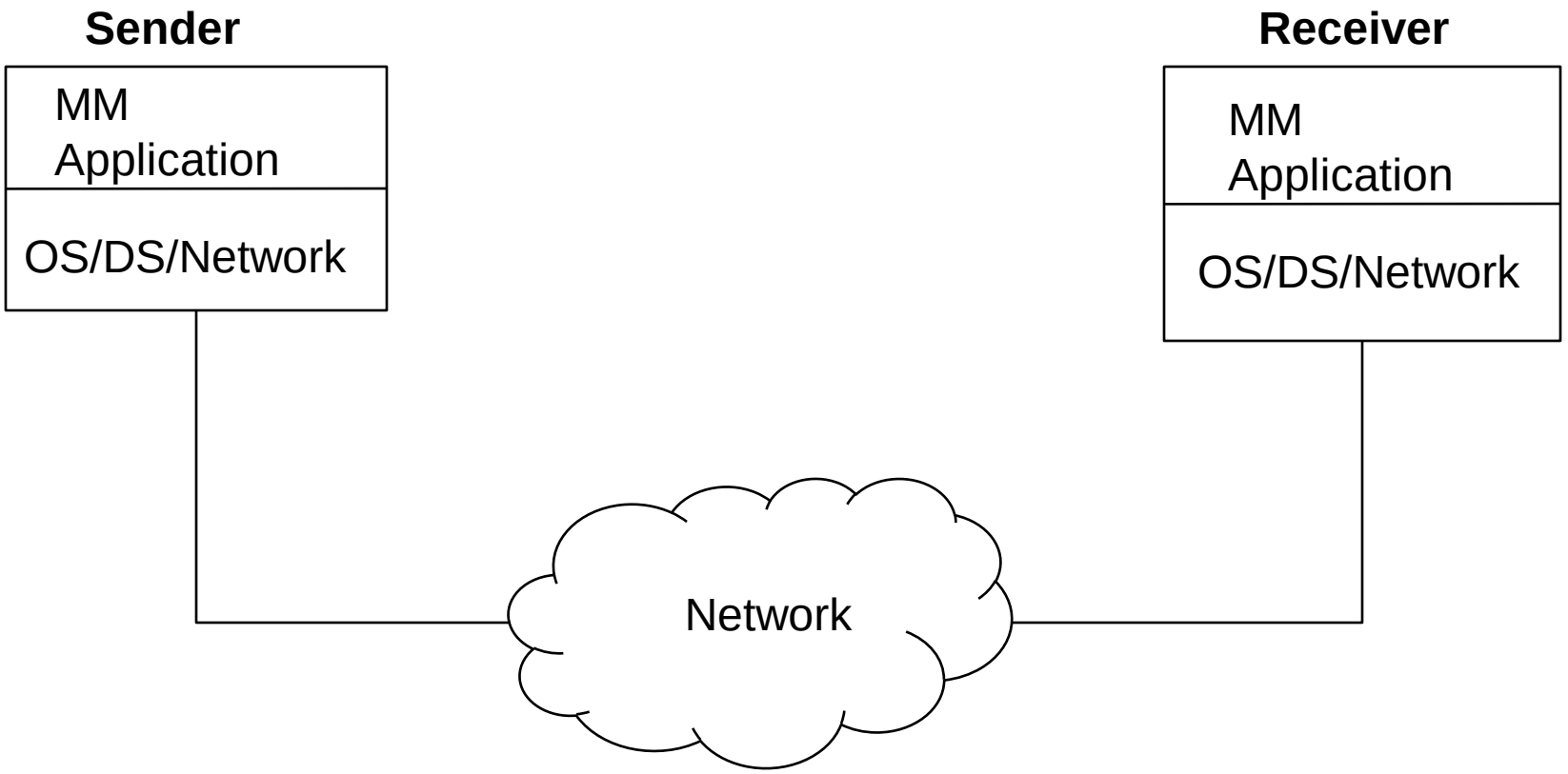
Klara Nahrstedt

Spring 2009

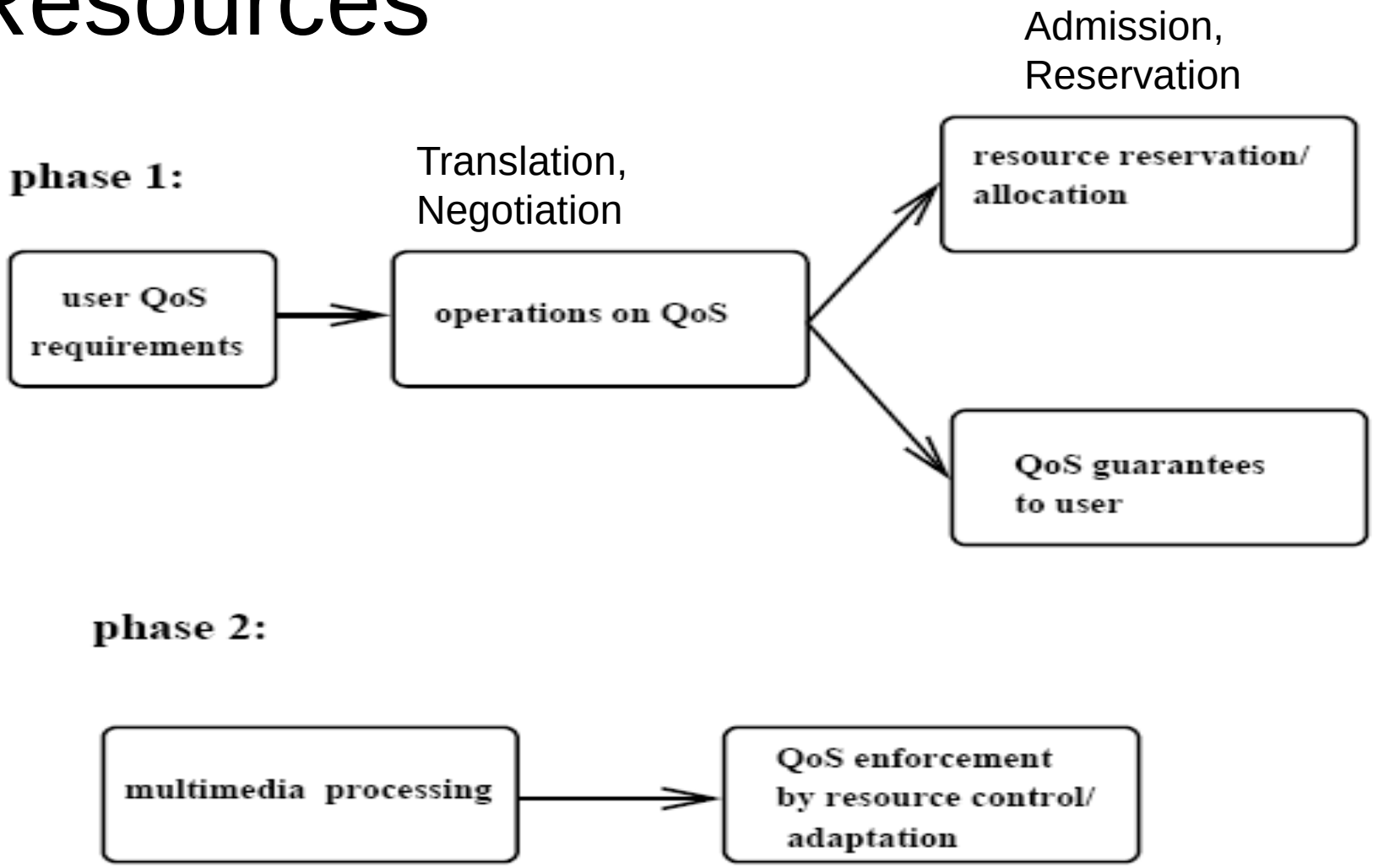
Administrative

- MP2 posted

Multimedia System/Network



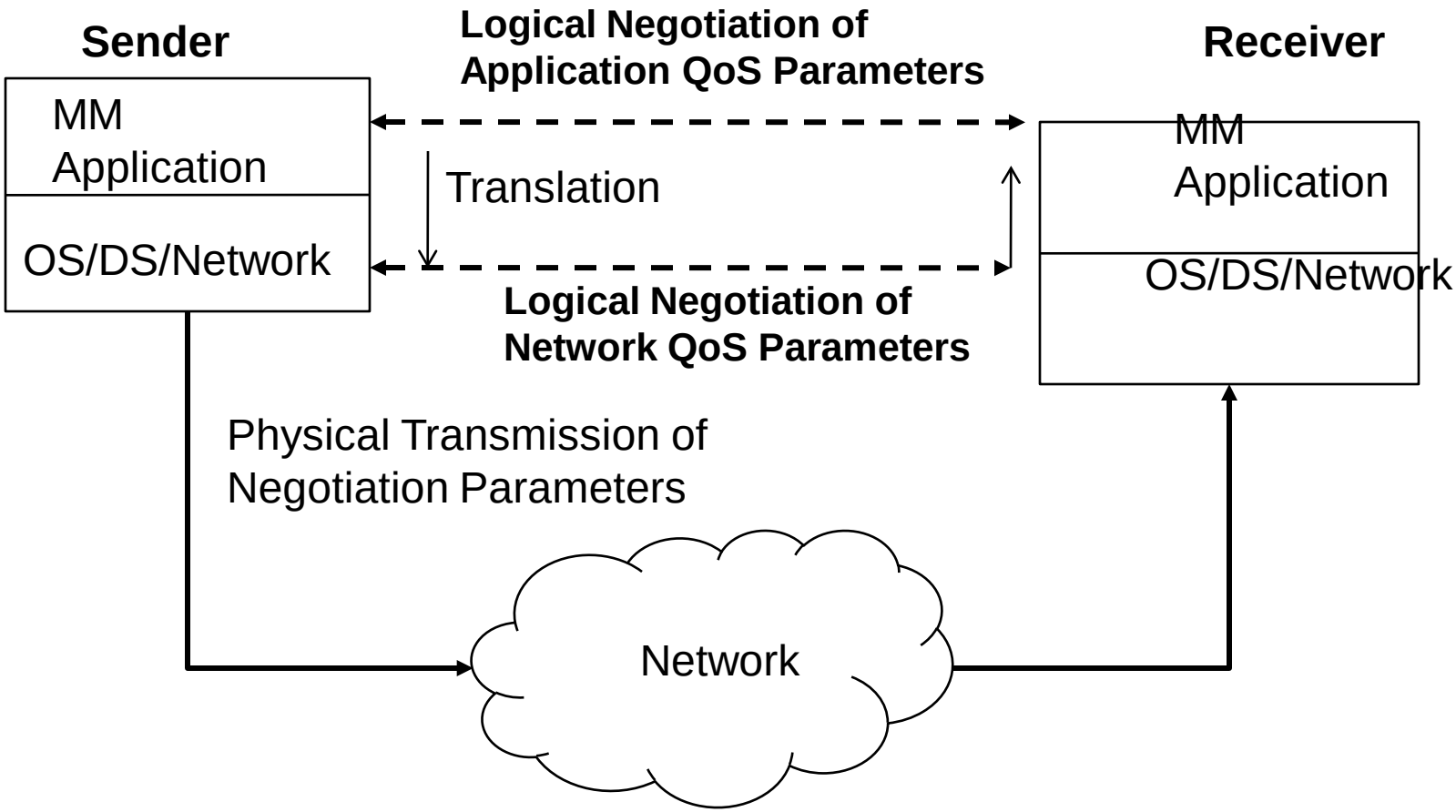
Relation between QoS and Resources



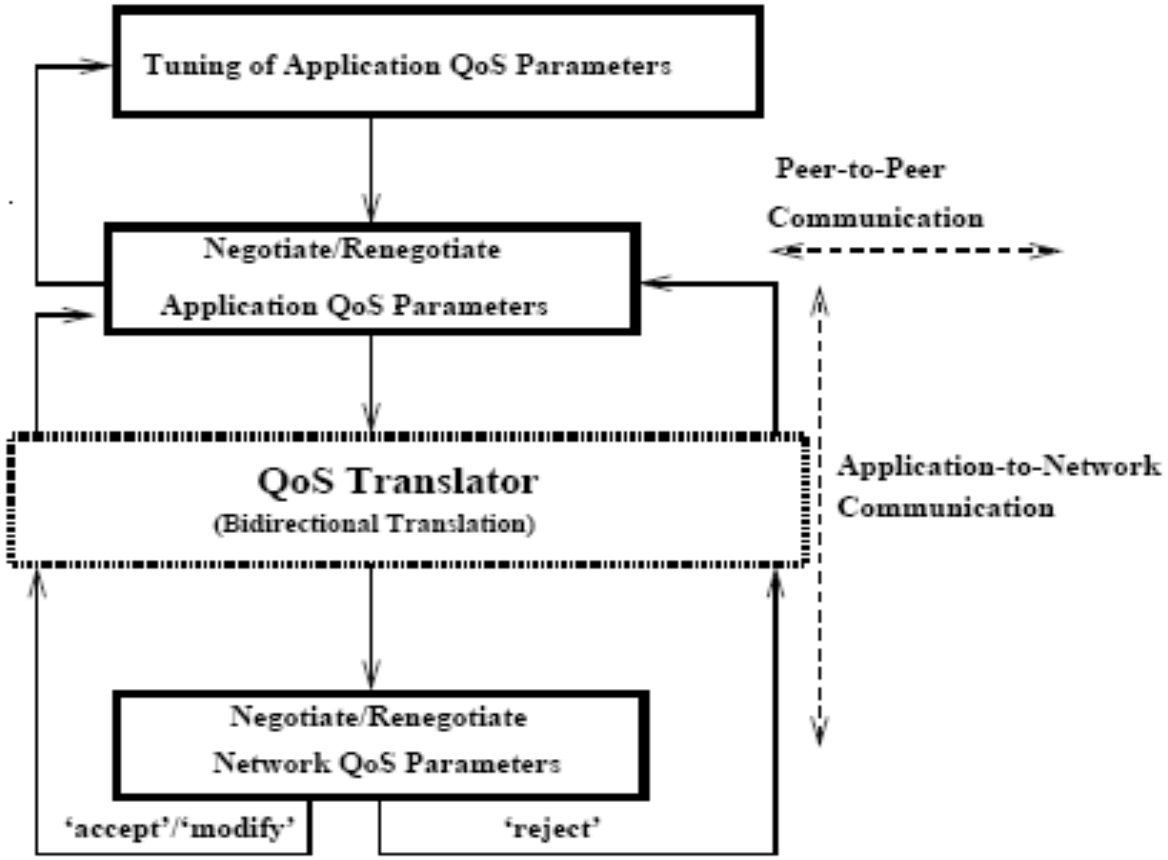
Phase 1: Establishment Phase (QoS Operations)

- **QoS Translation** at different Layers
  - User-Application
  - Application-OS/Transport Subsystem
- **QoS Negotiation**
  - Negotiation of QoS parameters among two peers/components

Phase 1: Connection Establishment



QoS Operations within Establishment Phase



User/Application
QoS Translation

Overlay P2P
QoS Negotiation

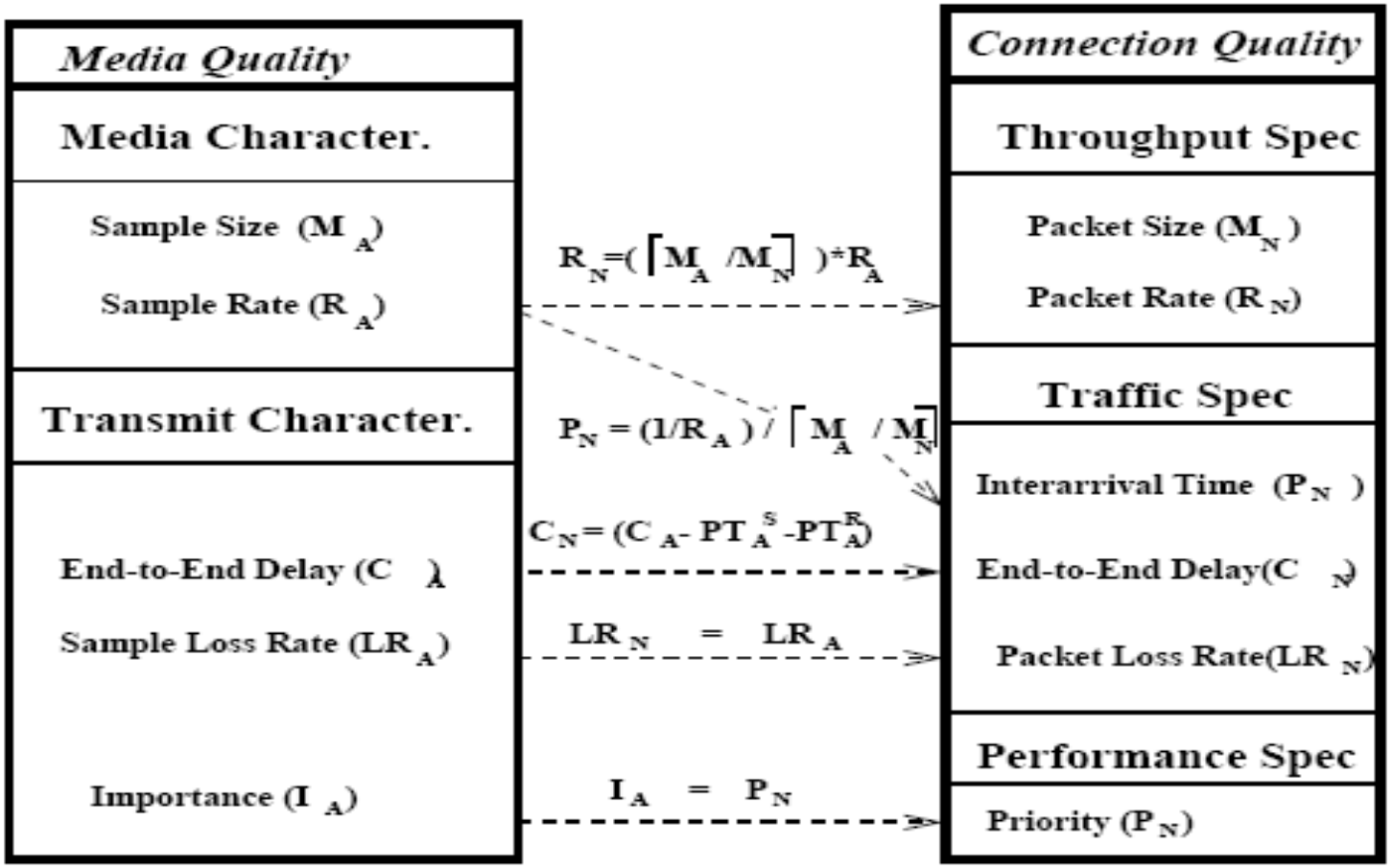
Application/Transport
QoS Translation

QoS Negotiation in
Transport Subsystem

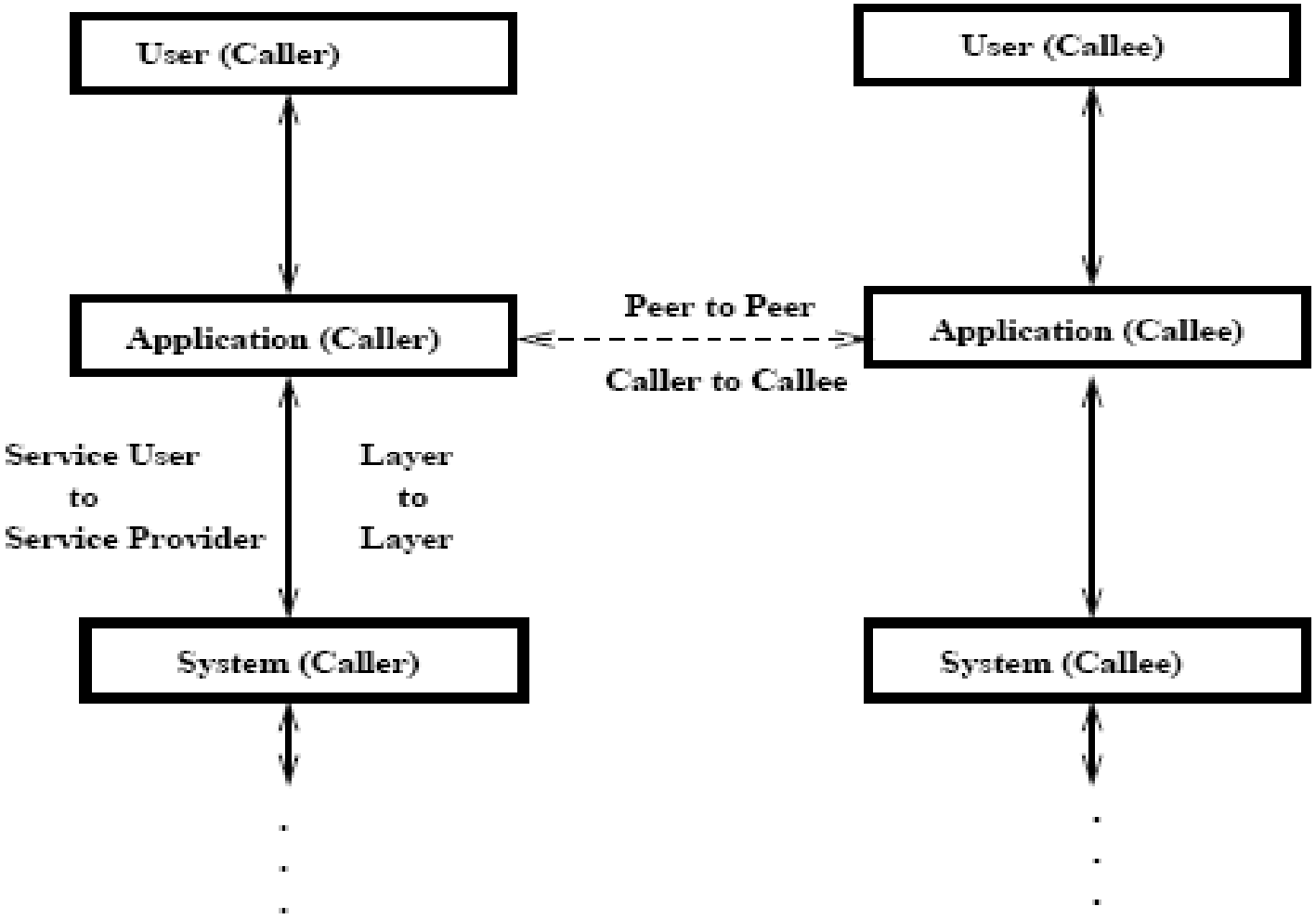
Example

- Video Stream Quality:
  - Frame size: 320x240 pixels, 24 bits (3 Bytes per pixel)
  - Application frame rate RA: 20 fps
- Translate to Network QoS if
  - Assume network packet size is 4KBytes
  - Network packet rate (RN):= ⌈320x240x3⌉ bytes / 4096 bytes

Layered Translation (Example)



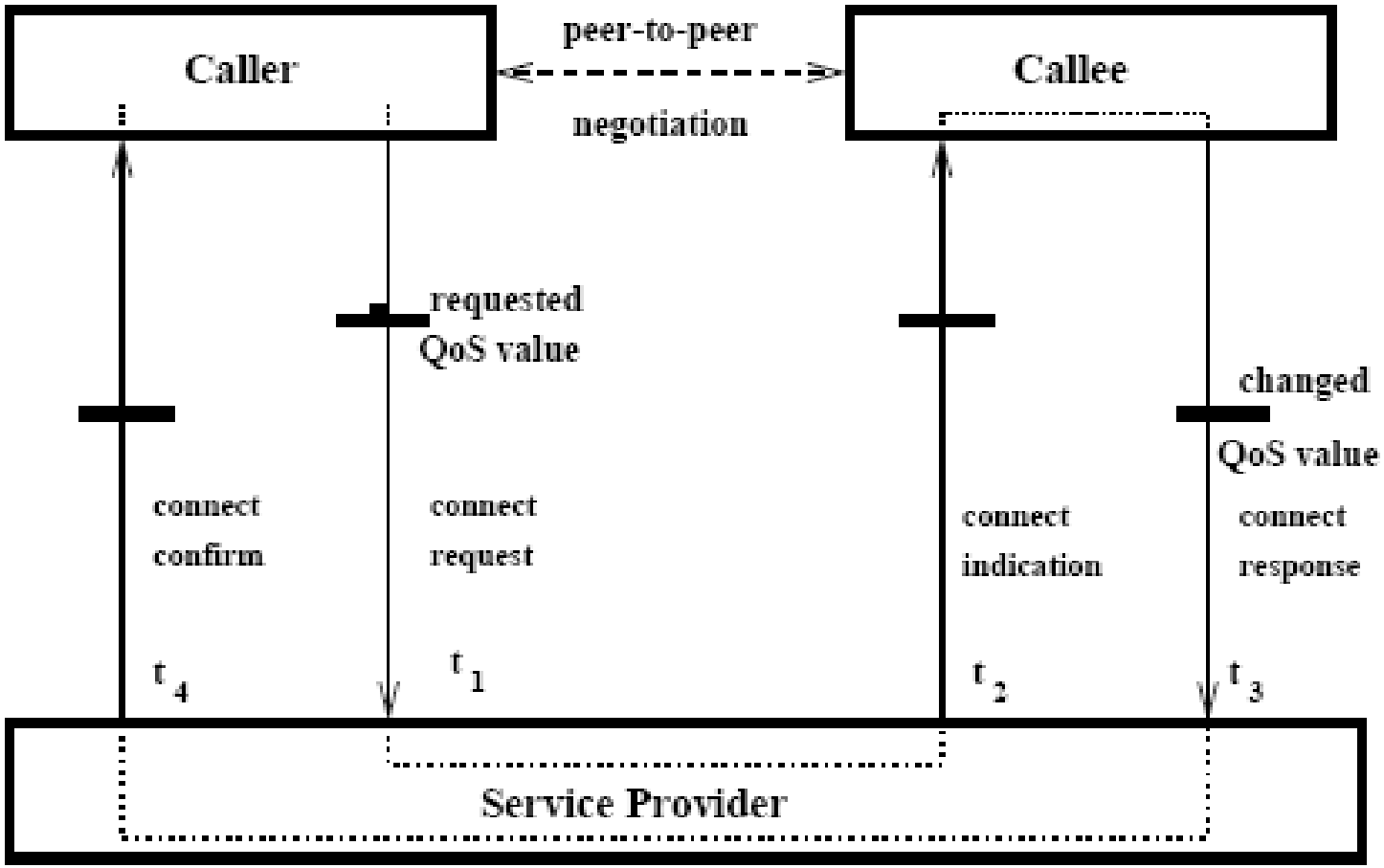
QoS Negotiation



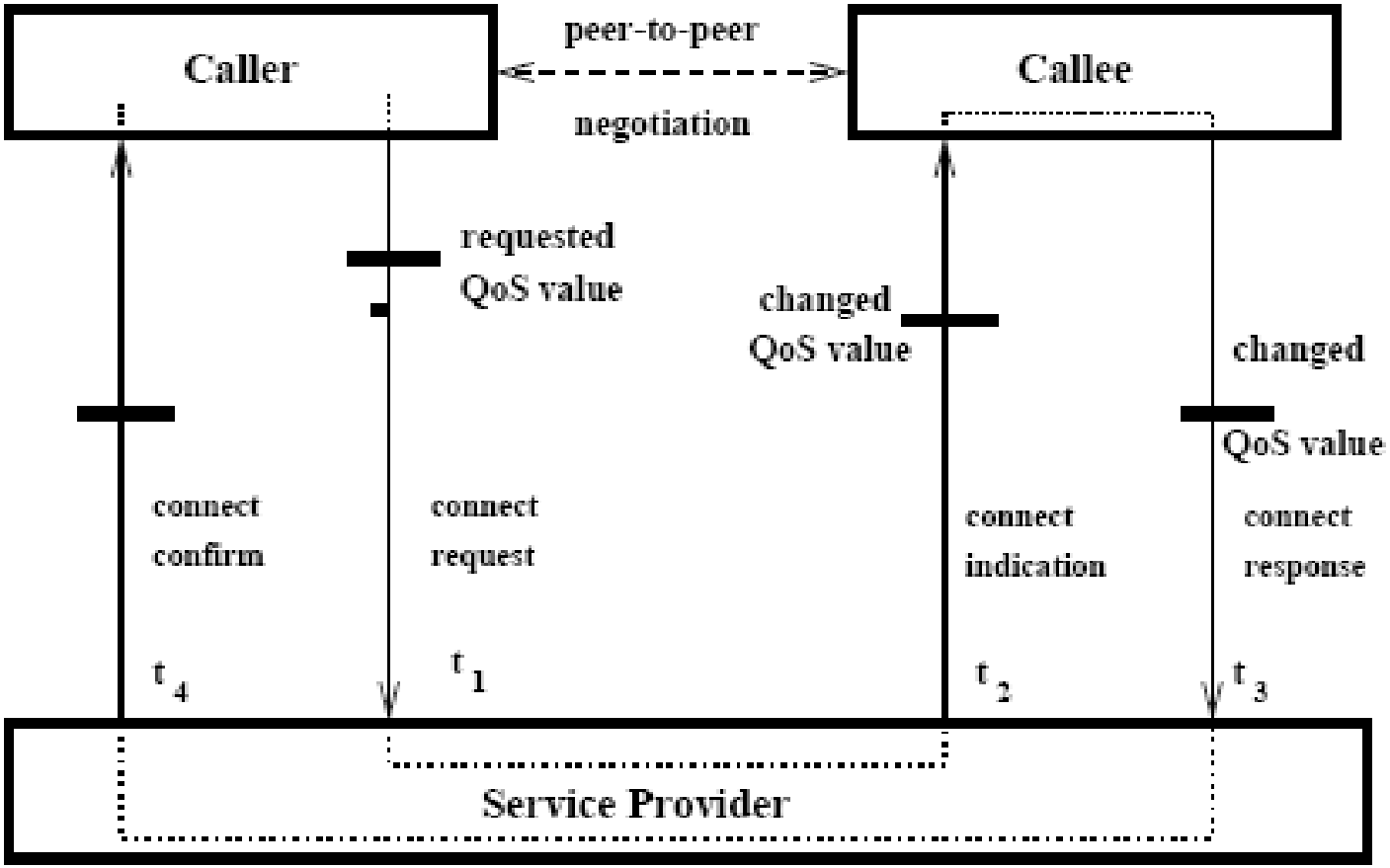
Different Types of Negotiation Protocols

- Bilateral Peer-to-Peer Negotiation
  - Negotiation of QoS parameters between equal peers in the same layer
- Triangular Negotiation
  - Negotiation of QoS parameters between layers
- Triangular Negotiation with Bounded Value

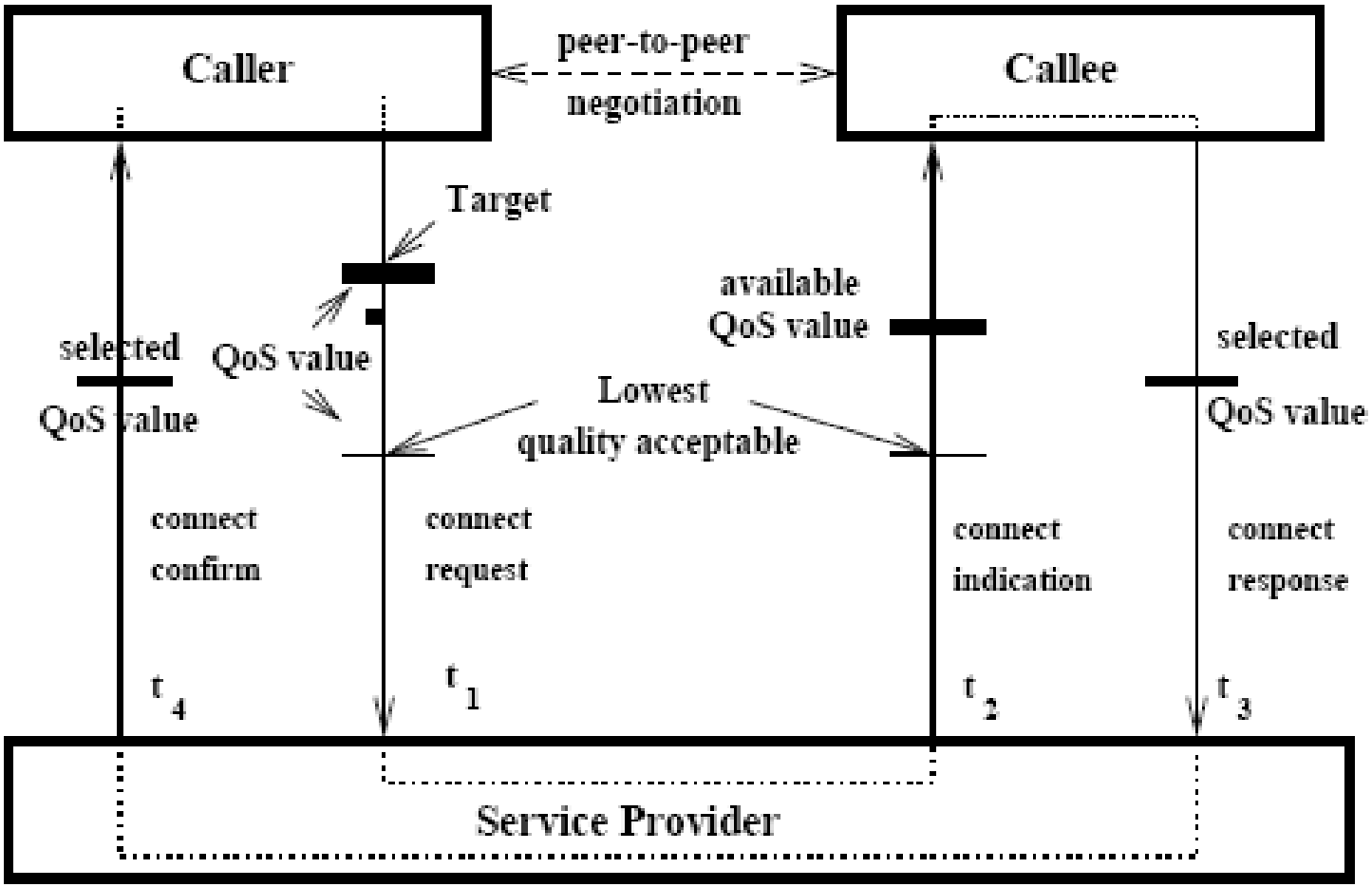
Bilateral QoS Negotiation



Triangular QoS Negotiation



Triangular Negotiation with Bounded Value



Multimedia Resource Management

- Resource managers with operations and resource management protocols
  - Various operations must be performed by resource managers in order to provide QoS
- Phase 1: Establishment Phase (resource operations)
  - Operations are executed where schedulable units utilizing shared resources must be admitted, reserved and allocated according to QoS requirements
- Phase 2: Enforcement Phase
  - Operations are executed where reservations and allocations must be enforced, and adapted if needed

Phase 1: Resource Preparation Operations

- QoS to Resource Mapping
  - Need translation or profiling (e.g., how much processing CPU cycles, i.e., processing time, it takes to process 320x240 pixel video frame)
- Resource Admission
  - Need admission tests to check availability of shared resources
- Resource Reservation
  - Need reservation mechanisms along the end-to-end path to keep information about reservations
- Resource Allocation

Continuous Media Resource Model

- One possible resource utilization model for multimedia data – Linear Bounded Arrival Process Model (LBAP)
- LBAP models message arrival process:
  - M – maximum message size (in bytes)
  - R – maximum message rate in messages per second
  - B – maximum burstiness (accumulation of messages)

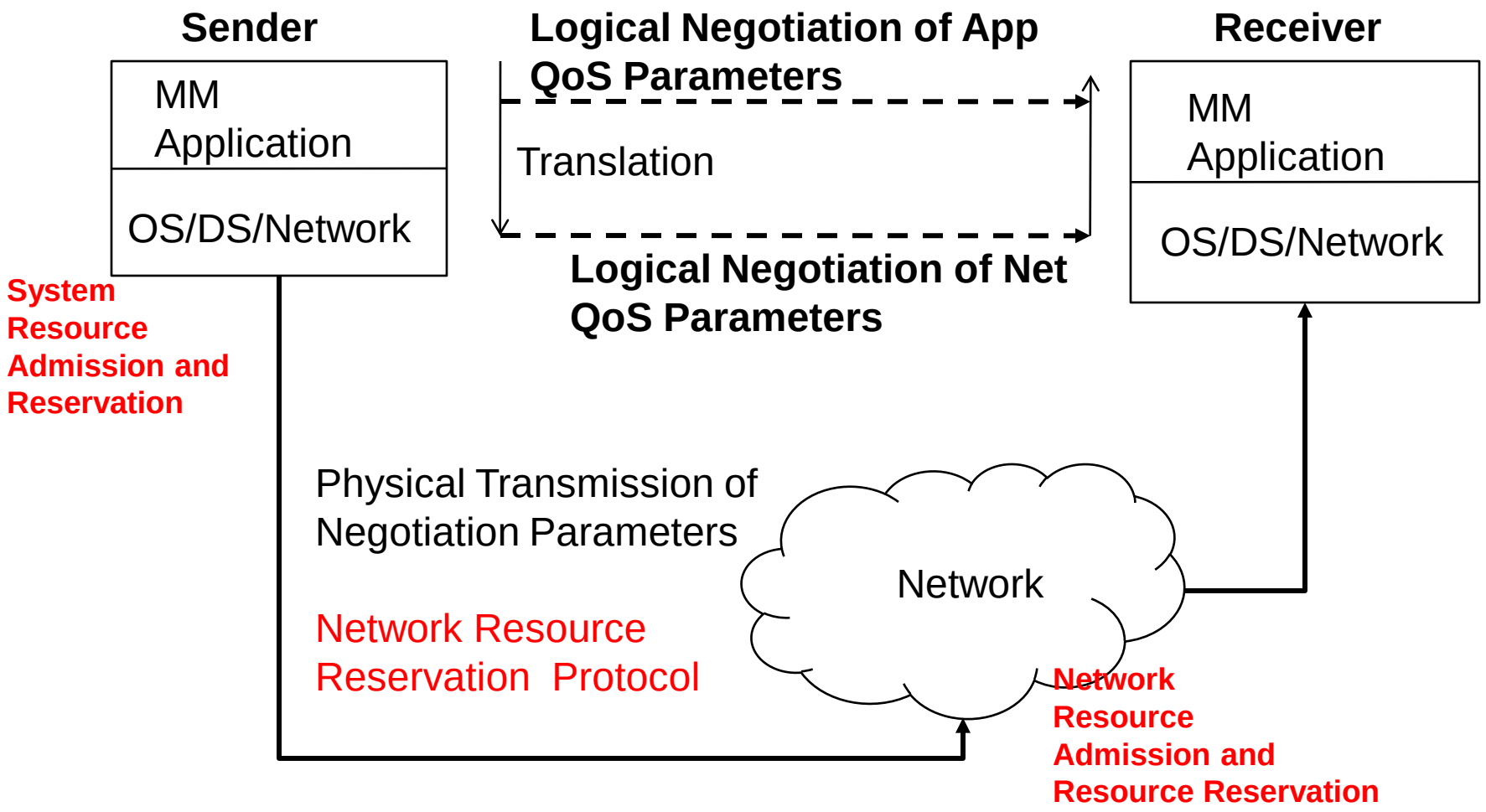
LBAP Resource Model

- If we have (M,R,B), we can predict utilization of resources:
  - Maximum number *N* of messages arriving at the resource: $N = B + R \times TimeInterval$
    - *Important for memory and CPU allocation*
  - Maximal Average Rate *R'* (in bytes per second): $R' = M \times R$
    - *Important for network bandwidth allocation*
  - Maximal Buffer Size (BS in bytes): $BS = M \times (B+1)$

Example of LBAP

- Consider M = 1176 Bytes per message, R = 75 messages per second, B = 10 messages
- During a time interval of 1 second, the maximum number of messages arriving at a resource must not exceed N = 10 messages + (75 messages/second * 1 second) = 85 messages
- Maximum average data rate in bytes per second is R' = 1176 bytes * 10 messages/second = 88200 bytes/second
- Maximum buffer size in bytes in BS = 1176 bytes * (10 messages + 1) = 12936 bytes

Phase 1: Connection Establishment



Admission Tests

- Task (System) schedulability tests for CPU resources
  - This is done for delay guarantees
- Network Packet schedulability tests for sharing host network interfaces, network switches
  - This is done for network delay and jitter guarantees
- Spatial tests for memory/buffer allocation
  - This is done for delay and reliability guarantees
- Network Link bandwidth tests
  - This is done for network throughput guarantees

Resource Reservation and Allocation

- Two types of reservations
  - Pessimistic approach - Worst case reservation of resources
  - Optimistic approach - Average case reservation of resources
- To implement resource reservation we need:
  - Resource table
    - to capture information about managed table (e.g., process management PID table)
  - Reservation table
    - to capture reservation information
  - Reservation function
    - to map QoS to resources and operate over reservation table

Resource Reservation

- Two types of reservation styles:
  - Sender-initiated reservation
  - Receiver-initiated reservation

Conclusion – Current State of Art

- Lack of mechanisms to support QoS guarantees
  - Need research in distributed control, monitoring, adaptation and maintenance of QoS mechanisms
- Lack of overall frameworks
  - Need QoS frameworks for heterogeneous environments (diverse networks, diverse devices, diverse OS)