

Κατανεμημένα Συστήματα

Μάθημα #5



Περιεχόμενα

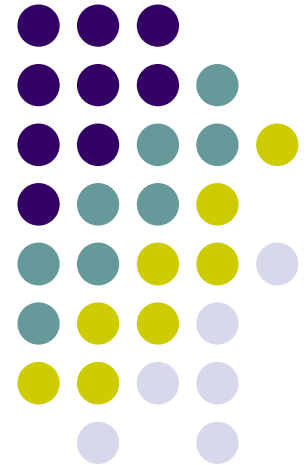


- Μηνυματοστρεφής Επικοινωνία
- Σύντομη αναφορά στα Berkeley Sockets
- Σύντομη αναφορά στο MPI
- Μηνυματοστρεφής διατηρούμενη επικοινωνία

Συστήματα Ουρών Μηνυμάτων

- IBM MQSeries (σύντομη αναφορά)
- Ρευματοστρεφής επικοινωνία

Μηνυματοστρεφής Επικοινωνία (Message-Oriented Communication)





Μηνυματοστρεφής Επικοινωνία

RPC/RMI are also known as request/response.

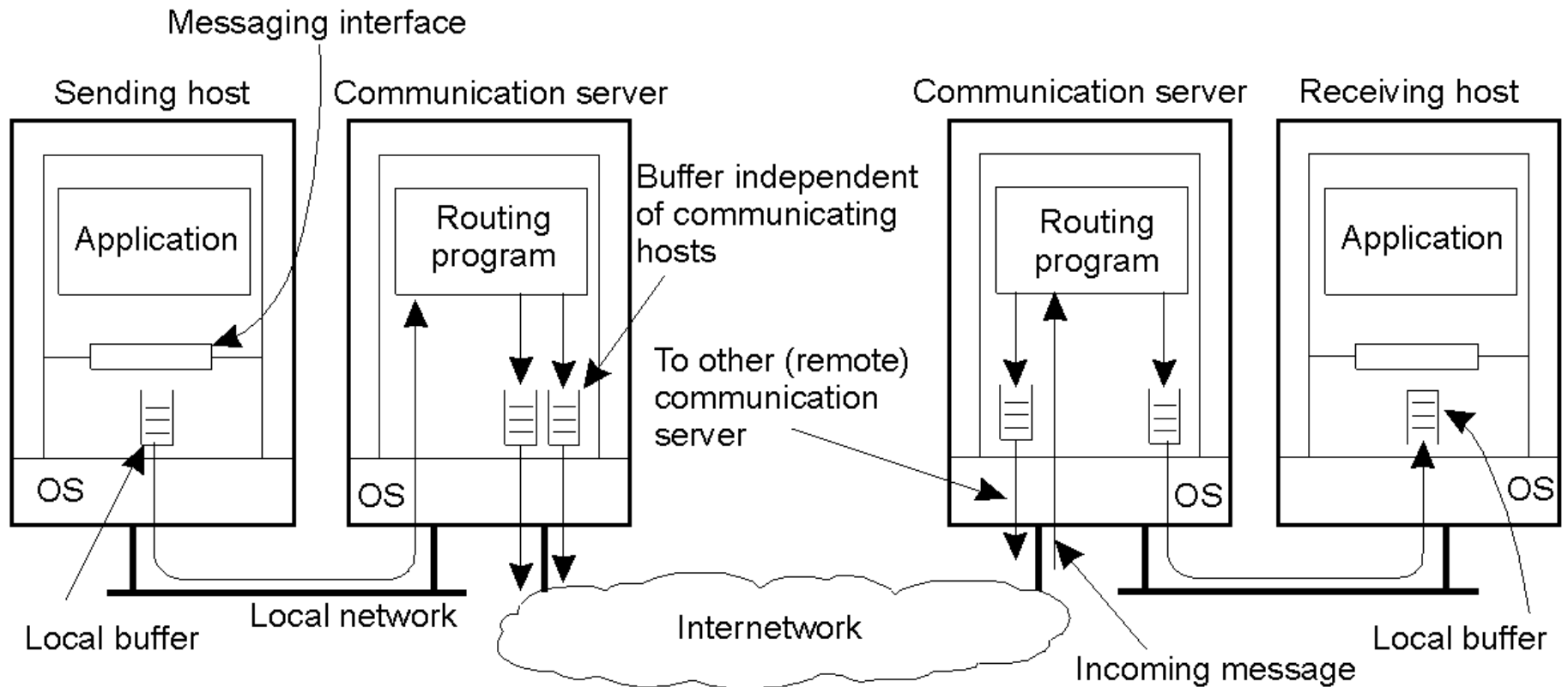
As a communications mechanism, RPC/RMI is often inappropriate.

For example: what happens if we cannot assume that the receiving side is “awake” and waiting to communicate?

The default “synchronous, blocking” nature of RPC/RMI is often too restrictive.

Something else is needed: Messaging.

Γενική οργάνωση συστήματος επικοινωνίας όπου οι υπολογιστές είναι συνδεδεμένοι μέσω δικτύου



Διατηρούμενη και Παροδική Επικοινωνία



Διατηρούμενη Επικοινωνία - Persistent Communication:

Μόλις στείλει το μήνυμα, ο αποστολέας μπορεί να σταματήσει. Ο παραλήπτης δεν χρειάζεται να είναι σε λειτουργία - το σύστημα αποθηκεύει το μήνυμα μέχρι να παραδοθεί.

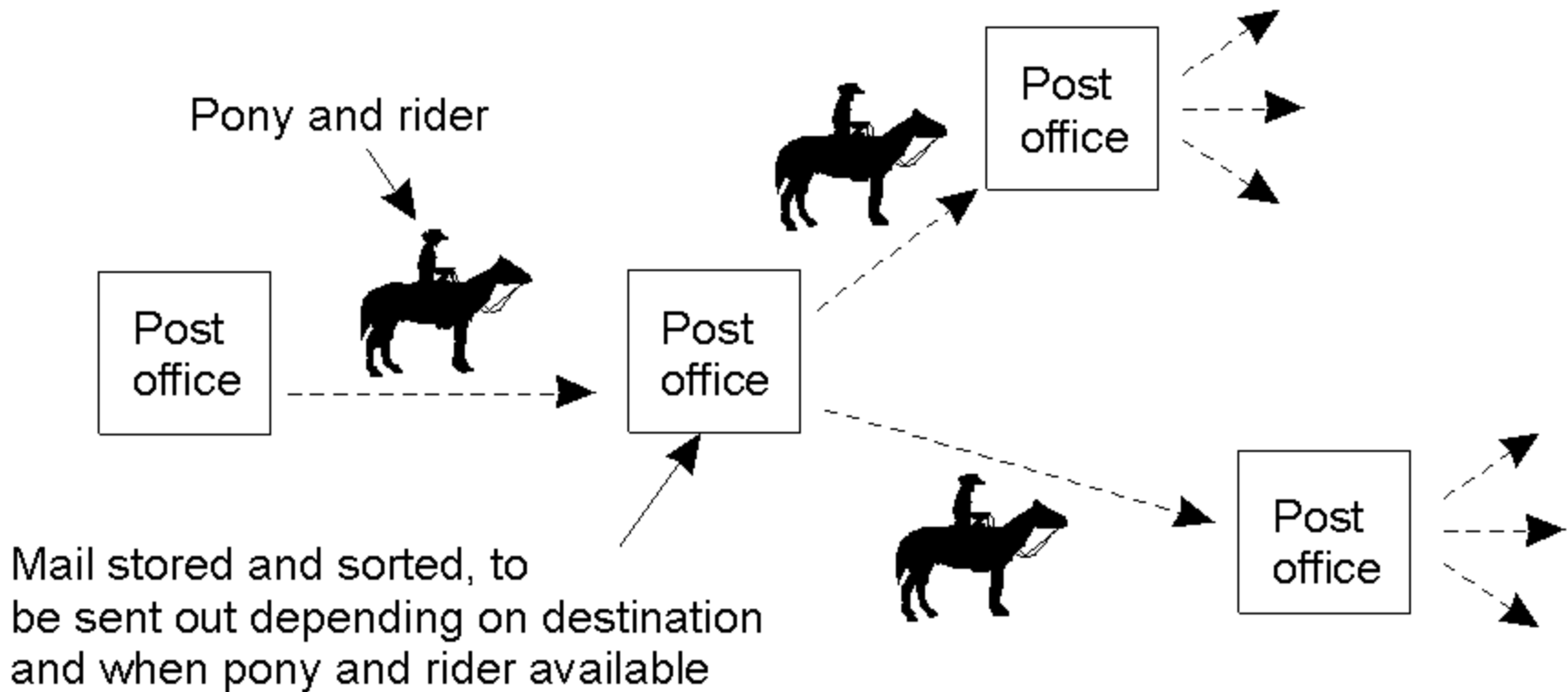
Παροδική Επικοινωνία - Transient Communication:

Το μήνυμα αποθηκεύεται μόνο για όσο διάστημα ο αποστολέας και ο παραλήπτης είναι σε λειτουργία. Εάν προκύψουν προβλήματα, το μήνυμα απορρίπτεται

**With transient communication, discard the message if next hop is down.
Persistent communication will store.**

- **How long should it store the message?**
- **There is also a spectrum between transient and persistent.**

Διατηρούμενη Επικοινωνία



- Persistent communication of letters back in the days of the Pony Express.
- Also known as **store-and-forward**.

Ασύγχρονη και Σύγχρονη Επικοινωνία



Ασύγχρονη Επικοινωνία :

Ο αποστολέας συνεχίζει με άλλες εργασίες αμέσως μετά την αποστολή του μηνύματος στον παραλήπτη.

Σύγχρονη Επικοινωνία :

Ο αποστολέας περιμένει απάντηση από τον παραλήπτη πριν κάνει οποιαδήποτε άλλη εργασία. (Αυτό τείνει να είναι το default μοντέλο για τις τεχνολογίες RPC /RMI).

Παροδική/Διατηρούμενη Επικοινωνία



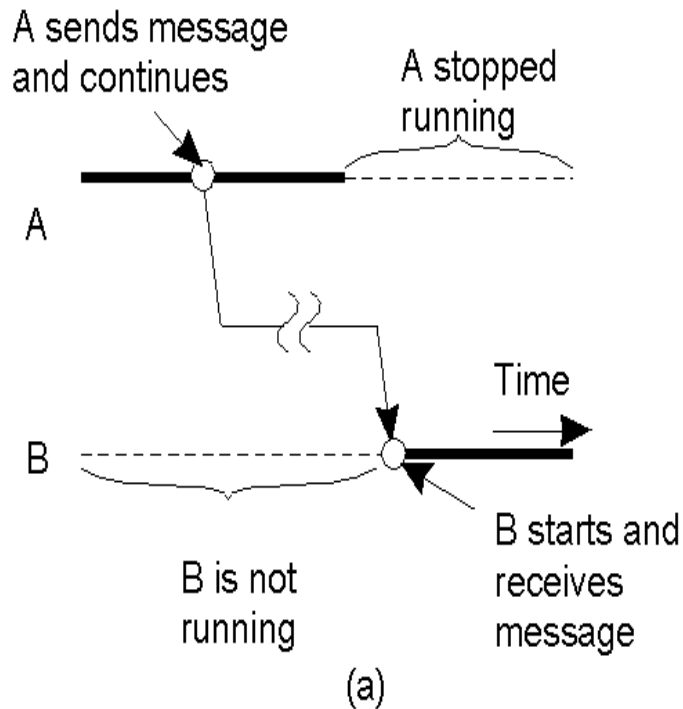
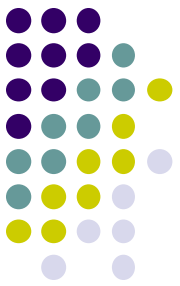
Παροδική επικοινωνία:

- Σε δίκτυα μεγάλης κλίμακας και ευρείας γεωγραφικής διασποράς η πρόσβαση μπορεί να είναι προβληματική λόγω αστοχιών
- Όταν παρουσιάζεται αστοχία θα πρέπει να συγκαλύπτεται αμέσως και να ξεκινά διαδικασία ανάκαμψης.

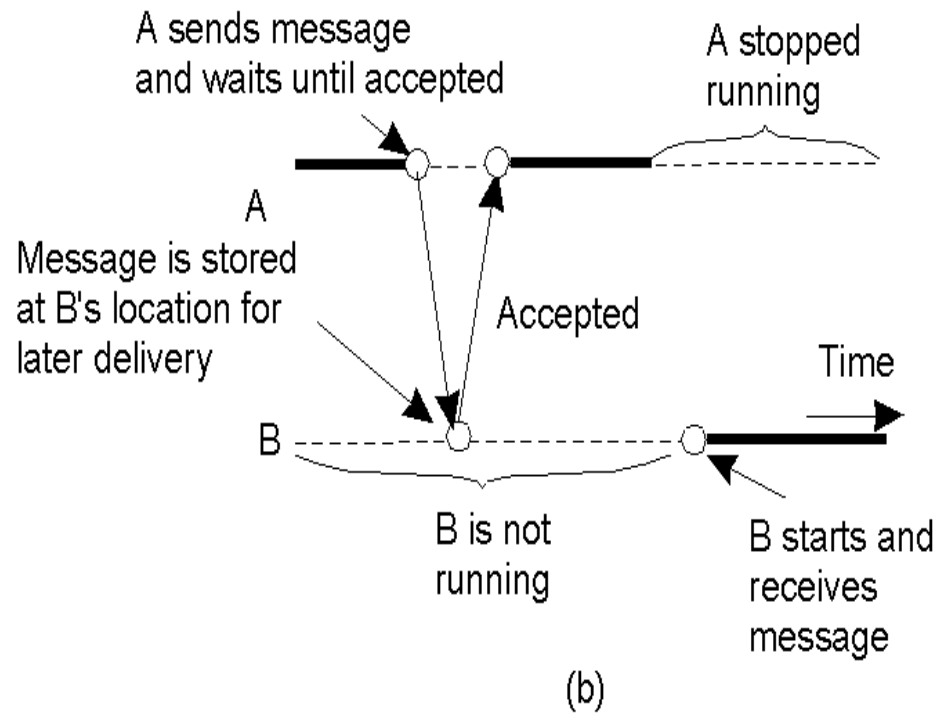
Διατηρούμενη επικοινωνία:

- Οι εφαρμογές σχεδιάζονται έτσι ώστε να μπορούν να χειρίζονται μακρόχρονες καθυστερήσεις μεταξύ αποστολής αίτησης και λήψης απάντησης
- Συνεπώς, απλούστερες και πιο αργές λύσεις για τη συγκάλυψη αστοχιών και την ανάκαμψη είναι αρκετές

Μορφές Επικοινωνίας

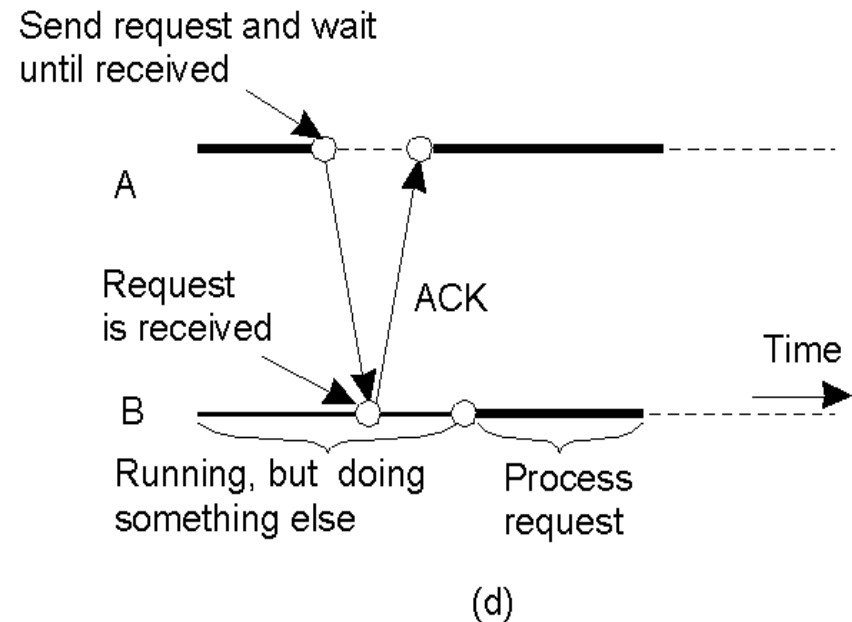
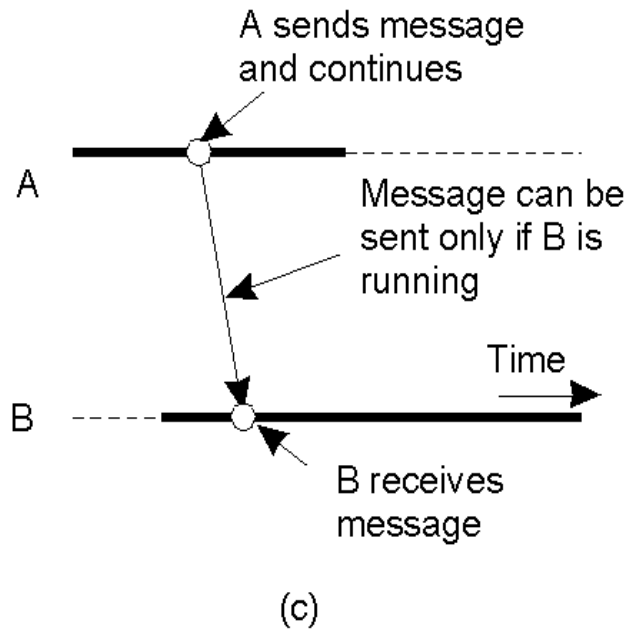
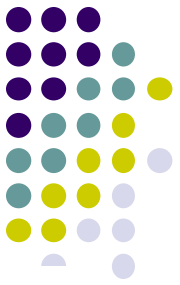


a) Διατηρούμενη ασύγχρονη επικοινωνία



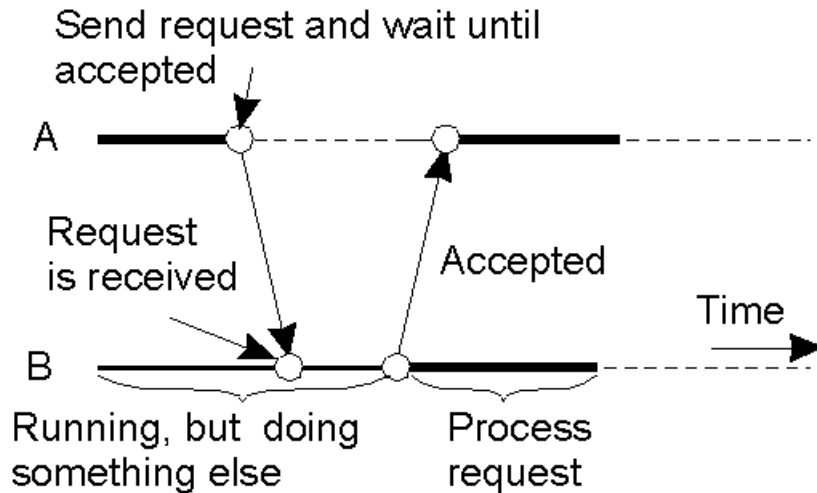
b) Διατηρούμενη σύγχρονη επικοινωνία

Μορφές Επικοινωνίας

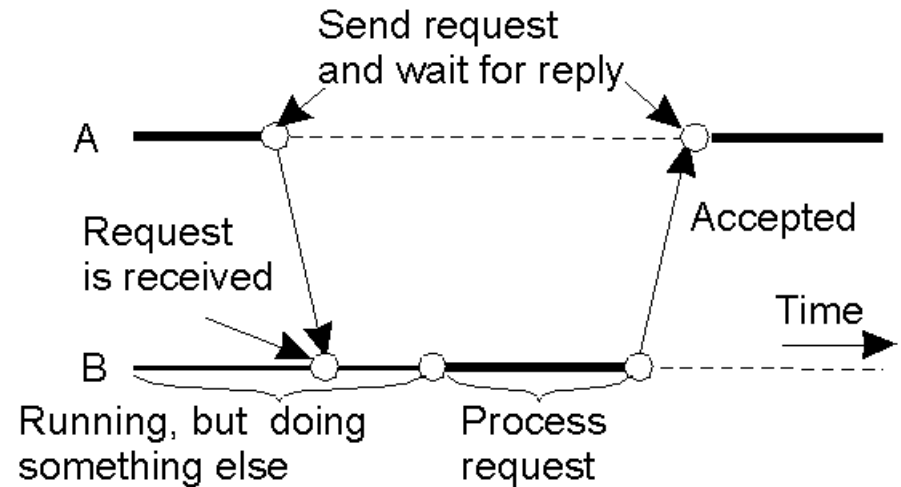


- c) Παροδική ασύγχρονη επικοινωνία
- d) Παροδική σύγχρονη επικοινωνία βασισμένη σε απόδειξη παραλαβής (ο αποστολέας μπλοκάρεται έως ότου το μήνυμα αποθηκευθεί σε χώρο προσωρινής αποθήκευσης στον παραλήπτη)

Μορφές Επικοινωνίας



(e)



(f)

- e) Παροδική σύγχρονη επικοινωνία βασισμένη στην παράδοση του μηνύματος (ο αποστολέας μπλοκάρεται έως ότου το μήνυμα επιδοθεί στον παραλήπτη για περαιτέρω επεξεργασία)
- f) Παροδική σύγχρονη επικοινωνία βασισμένη στην απόκριση

Μηνυματοστρεφής Παροδική Επικοινωνία



- Υποδοχές Berkeley (Berkeley Sockets)
- The Message-Passing Interface (MPI)



Υποδοχές Berkeley (Berkeley Sockets)

Υποδοχή ονομάζεται ένα ακραίο σημείο επικοινωνίας στο οποίο γράφονται δεδομένα και από το οποίο διαβάζονται δεδομένα.

Μια υποδοχή αποτελεί ένα επίπεδο αφαίρεσης πάνω από το πραγματικό ακραίο σημείο επικοινωνίας που χρησιμοποιείται από το τοπικό ΛΣ για ένα συγκεκριμένο πρωτόκολλο μεταφοράς.



Υποδοχές Berkeley

Θεμελιώδεις λειτουργίες υποδοχών για το TCP/IP.

Primitive	Meaning
Socket	Create a new communication endpoint
Bind	Attach a local address to a socket
Listen	Announce willingness to accept connections
Accept	Block caller until a connection request arrives
Connect	Actively attempt to establish a connection
Send	Send some data over the connection
Receive	Receive some data over the connection
Close	Release the connection

Θεμελιώδεις λειτουργίες υποδοχών για το TCP/IP



- Ο διακομιστής εκτελεί τις 4 πρώτες λειτουργίες με τη συγκεκριμένη σειρά.
- Η καλούσα διεργασία μέσω της socket δημιουργεί ένα νέο ακραίο σημείο (το ΛΣ δεσμεύει πόρους για την αποστολή/παραλαβή μηνυμάτων για το καθορισμένο πρωτόκολλο)
 - Η bind συσχετίζει μια τοπική διεύθυνση (IP διεύθυνση και θύρα) σε μια υποδοχή (ο διακομιστής θέλει να λαμβάνει μηνύματα μόνο στην καθορισμένη διεύθυνση IP και θύρα)
 - Η listen καλείται μόνο στην συνδεσμική επικοινωνία και δεν προκαλεί μπλοκάρισμα (επιτρέπει στο ΛΣ να κρατήσει δεσμευμένο αποθηκευτικό χώρο για έναν καθορισμένο αριθμό συνδέσεων που είναι διατεθειμένη να δεχθεί η καλούσα διεργασία)
 - Η κλήση της accept μπλοκάρει την καλούσα διεργασία έως ότου να ληφθεί μια αίτηση σύνδεσης.

Θεμελιώδεις λειτουργίες υποδοχών για το TCP/IP

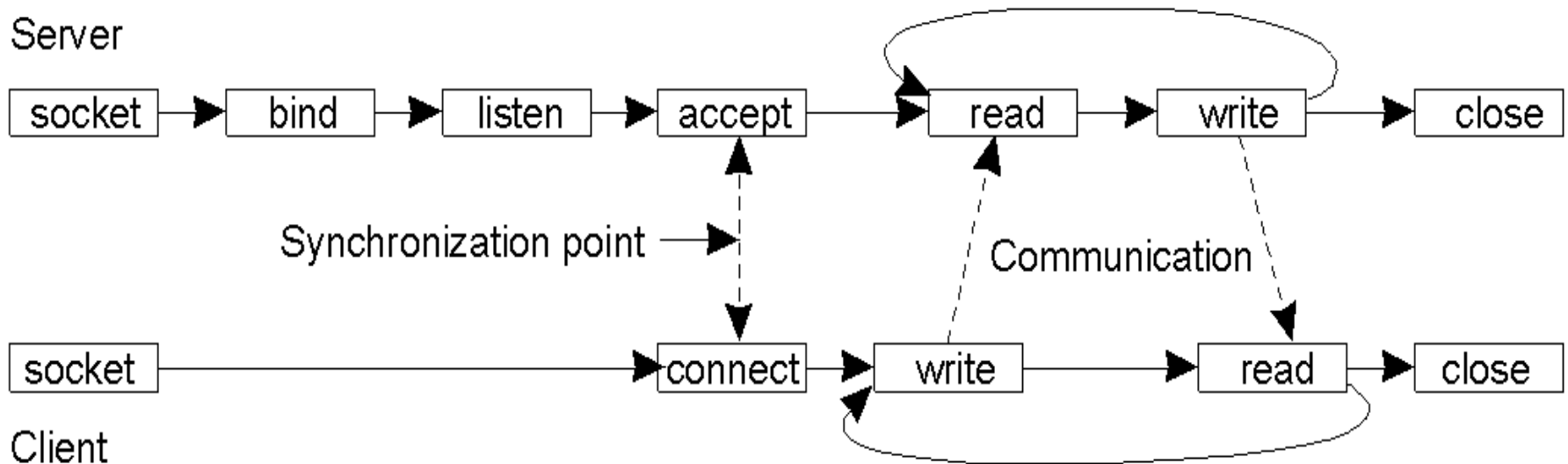


Στην πλευρά του πελάτη

- δημιουργείται υποδοχή με τη λειτουργία `socket`
- η κλήση της `bind` δεν είναι απαραίτητη αφού το ΛΣ εκχωρεί δυναμικά μια θύρα όταν εγκατασταθεί μια σύνδεση
- η `connect` απαιτεί από την καλούσα διεργασία τον προσδιορισμό της διεύθυνσης επιπέδου μεταφοράς όπου θα σταλεί η αίτηση σύνδεσης
- ο πελάτης μπλοκάρεται έως ότου εγκατασταθεί η σύνδεση

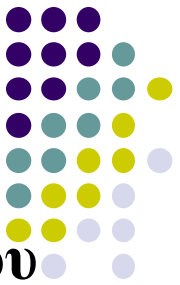
Όταν εγκατασταθεί η σύνδεση ανταλλάσσονται πληροφορίες μέσω των λειτουργιών `send` and `receive`.

Υποδοχές Berkeley



- Connection-oriented communication pattern using sockets.

Υποδοχές Berkeley



Ανάγκη για μηνυματοστρεφείς θεμελιώδεις λειτουργίες που επιτρέπουν την εύκολη συγγραφή εφαρμογών υψηλής απόδοσης

Οι υποδοχές είναι ανεπαρκείς διότι:

1. δεν παρέχουν το κατάλληλο επίπεδο αφαίρεσης αφού υποστηρίζουν μόνο θεμελιώδεις λειτουργίες αποστολής και λήψης
2. έχουν σχεδιασθεί για επικοινωνία με τη χρήση στοιβών πρωτοκόλλων γενικής χρήσης (π.χ. TCP/IP) και όχι με πρωτόκολλα δικτύων διασύνδεσης υψηλών ταχυτήτων τα οποία απαιτούν μια διασύνδεση που θα είναι σε θέση να χειριστεί πιο προηγμένες δυνατότητες (π.χ., διάφορες μορφές προσωρινής αποθήκευσης και συγχρονισμού).



Message Passing Interface (MPI)

- Sockets are too low level for scientific computing.
 - No data types.
 - No broadcast.
 - No collective operations.
 - No “message” abstraction.
- MPI was written to address that.
 - MPI job starts as a group of processes.
 - `mpirun prog_name host_list`
 - Each process has an integer “rank”, which serves as the address of that process. Nothing like an IP address.
 - MPI has specialized implementations over specialized hardware.
 - Infiniband
 - Myrinet
 - Quadrics
 - FibreChannel

The Message-Passing Interface (MPI)



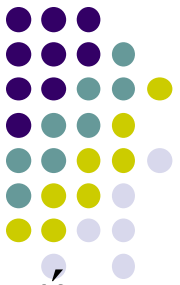
Some of the most intuitive message-passing primitives of MPI.

Primitive	Meaning
MPI_bsend	Append outgoing message to a local send buffer
MPI_send	Send a message and wait until copied to local or remote buffer
MPI_ssend	Send a message and wait until receipt starts
MPI_sendrecv	Send a message and wait for reply
MPI_issend	Pass reference to outgoing message, and continue
MPI_issend	Pass reference to outgoing message, and wait until receipt starts
MPI_recv	Receive a message; block if there are none
MPI_irecv	Check if there is an incoming message, but do not block



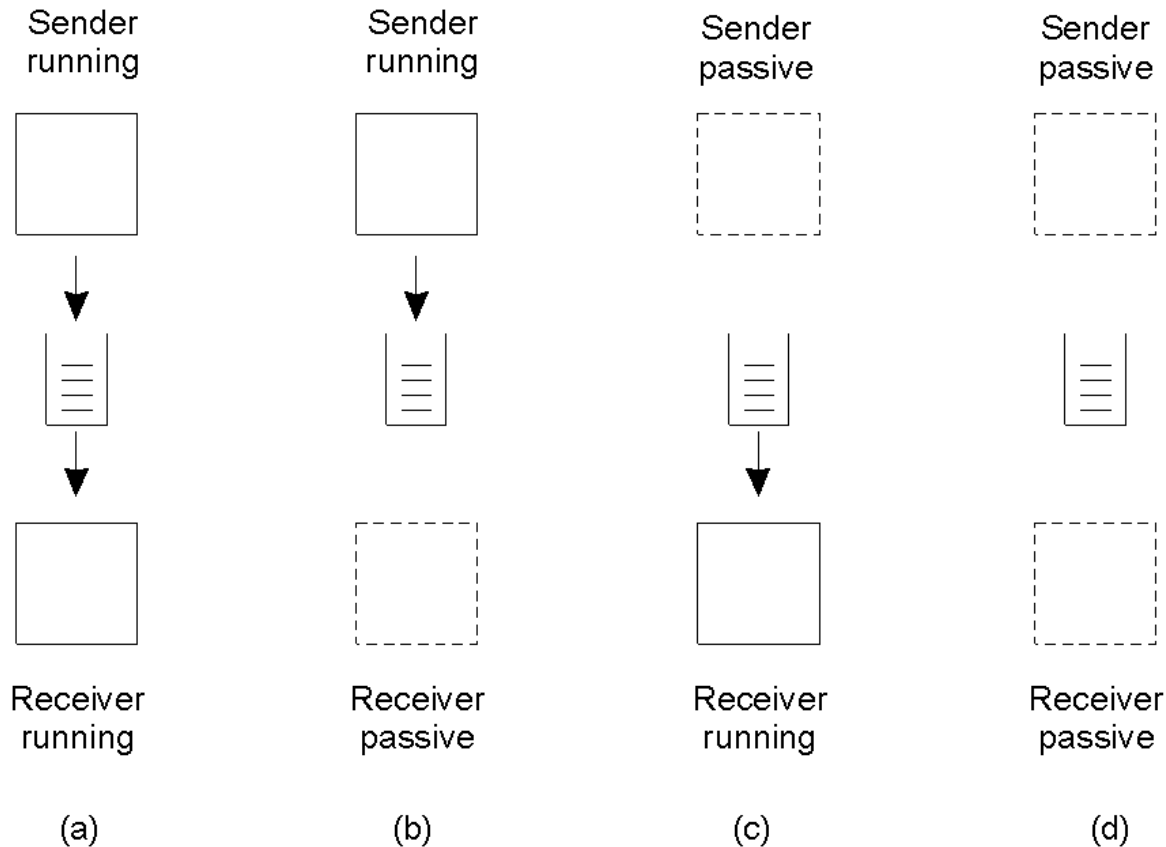
- MPI - sockets : υποστηρίζουν κυρίως παροδική επικοινωνία.

Συστήματα Ουρών Μηνυμάτων (Message-Queuing Systems)



- **Βασική ιδέα:** οι εφαρμογές επικοινωνούν με την εισαγωγή μηνυμάτων σε ουρές και την εξαγωγή τους από αυτές.
- Η μόνη εγγύηση που παρέχεται στον αποστολέα είναι ότι το μήνυμά του θα εισαχθεί στην ουρά του παραλήπτη
- Καλούνται και **συστήματα μηνυματοστρεφούς ενδιάμεσου λογισμικού (Message Oriented Middleware – MOM).**
- Υποστηρίζουν **διατηρούμενη ασύγχρονη επικοινωνία**
- Πραγματοποίηση χαλαρά συνδεδεμένης (“loosely - coupled”) επικοινωνίας. Ο αποστολέας και ο παραλήπτης εκτελούνται ανεξάρτητα ο ένας από τον άλλο.
- **Τέσσερις συνδυασμοί** σχετικά με την κατάσταση εκτέλεσης του αποστολέα και του παραλήπτη.

Συστήματα Ουρών Μηνυμάτων



Είδη χαλαρά συνδεδεμένης (“loosely -coupled”) επικοινωνίας
με τη χρήση ουρών



Συστήματα Ουρών Μηνυμάτων

Διευθυνσιοδότηση μηνύματος: γίνεται χρήση ενός ονόματος ουράς προορισμού το οποίο είναι μοναδικό για όλο το σύστημα.

Basic interface to a queue in a message-queuing system.

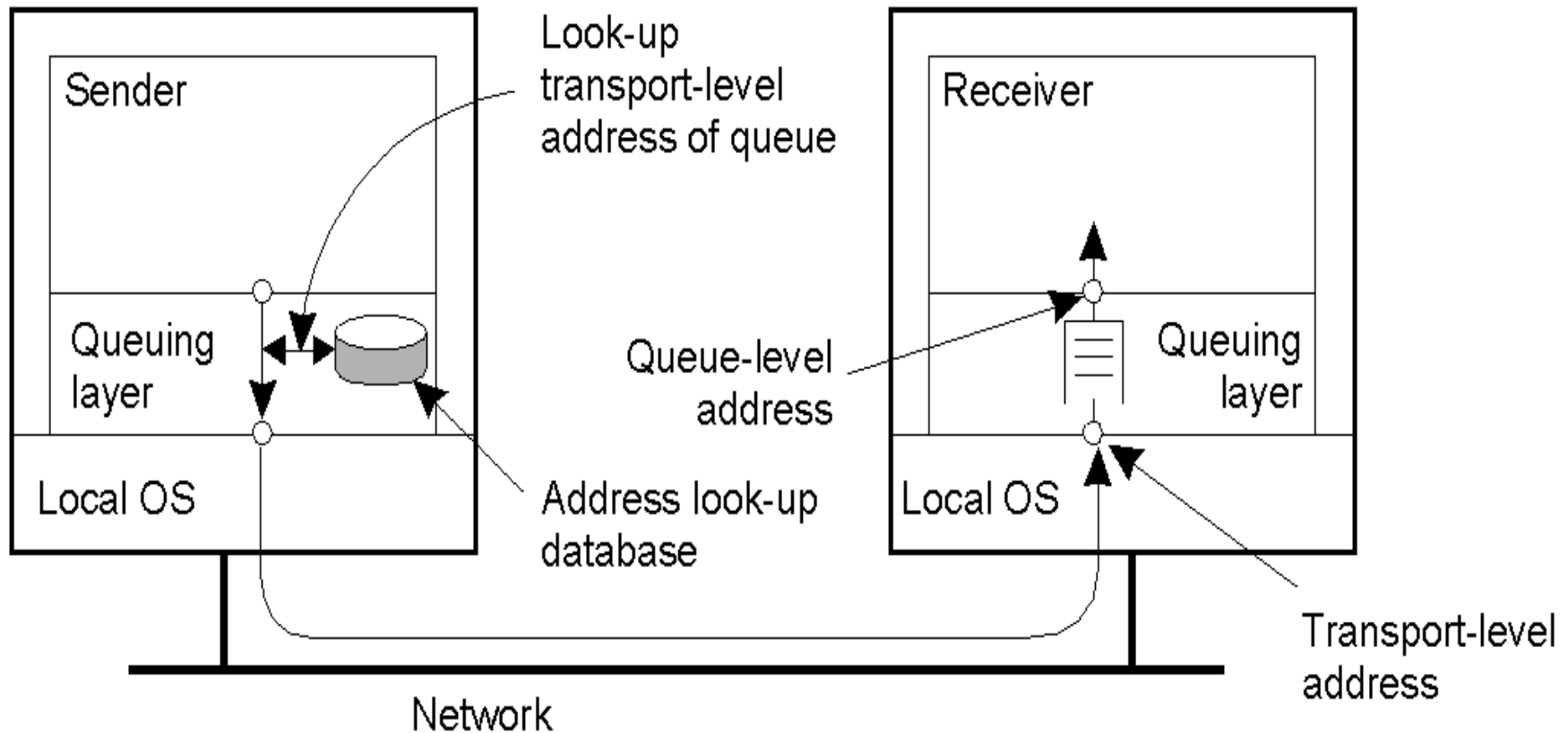
Primitive	Meaning
Put	Append a message to a specified queue
Get	Block until the specified queue is nonempty, and remove the first message
Poll	Check a specified queue for messages, and remove the first. Never block.
Notify	Install a handler to be called when a message is put into the specified queue.

Αρχιτεκτονική Συστήματος Ουρών Μηνυμάτων



- Κάθε μήνυμα τοποθετείται σε μια **ουρά αφετηρίας** (*source queue*) και εξάγεται από μια **ουρά προορισμού** (*destination queue*).
- Κάθε μήνυμα περιέχει τον προσδιορισμό μιας ουράς προορισμού στην οποία θα πρέπει να μεταφερθεί.
- Η διαχείριση των ουρών γίνεται με **διαχειριστές ουρών** (*Queue Managers*)
- Η μεταφορά είναι ευθύνη του συστήματος ουρών μηνυμάτων.
- Τηρείται βάση δεδομένων **αντιστοιχίας ουρών μηνυμάτων και θέσεων δικτύου**.

Σχέση μεταξύ διευθύνσεων επιπέδου ουρών και διευθύνσεων επιπέδου δικτύου

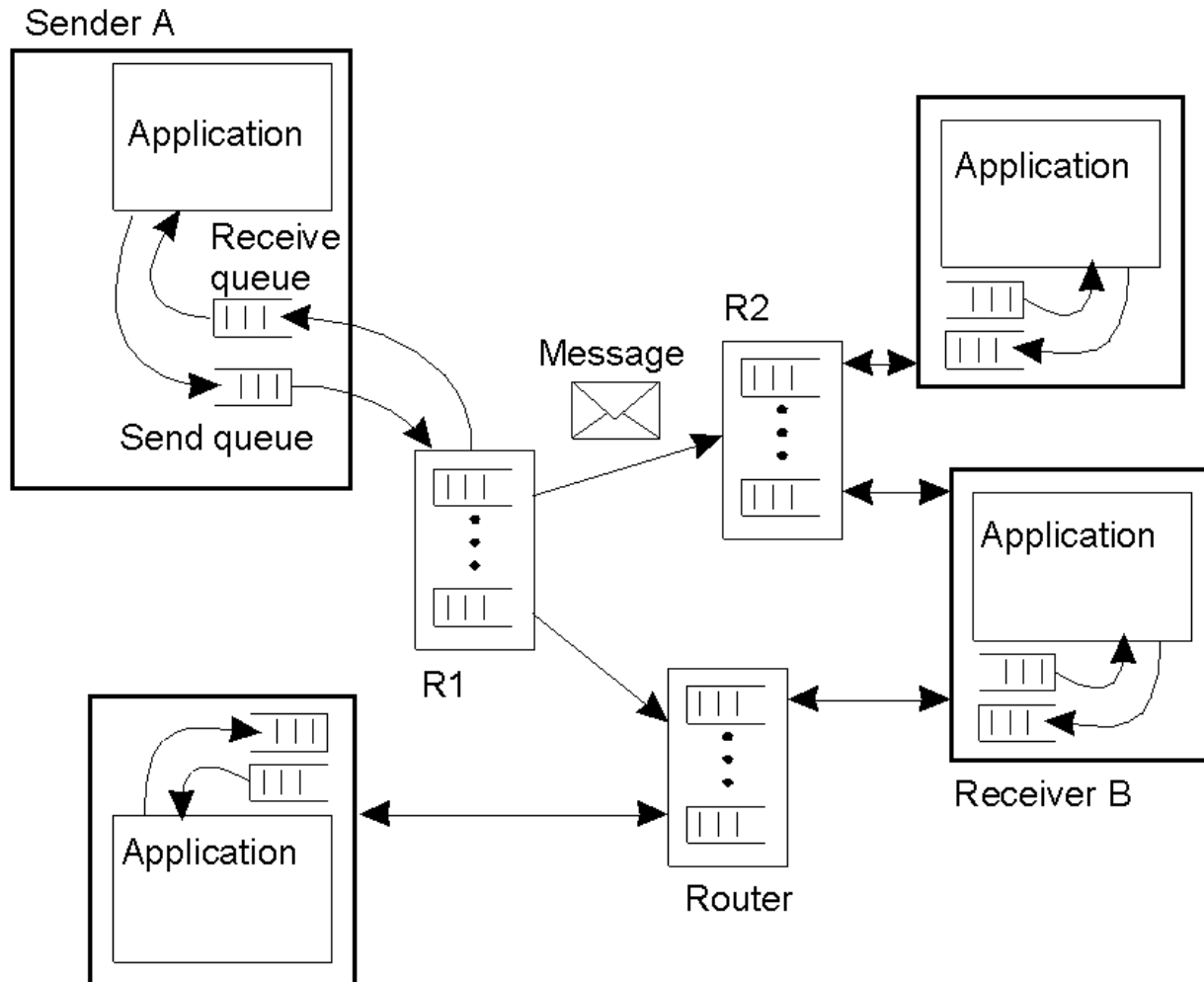
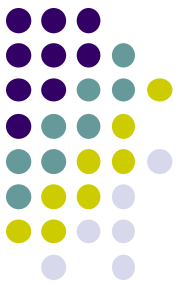


Οργάνωση ενός συστήματος ουρών μηνυμάτων με δρομολογητές



- Η διαχείριση των ουρών γίνεται με ειδικούς **διαχειριστές ουρών** που δρουν ως **δρομολογητές ή αναμεταδότες (relays)** δηλ., επικοινωνούν με τις εφαρμογές και προωθούν τα μηνύματα σε άλλους **δρομολογητές**.
- Έτσι τα συστήματα ουρών μηνυμάτων εξελίσσονται σε **δίκτυα επικάλυψης (“overlay networks”)** του επιπέδου εφαρμογής πάνω σε κάποιο δίκτυο υπολογιστών.
- Επιτυγχάνεται **επεκτασιμότητα**: όταν προστίθενται ή αφαιρούνται ουρές ενημερώνονται μόνον οι δρομολογητές ενώ οι υπόλοιποι διαχειριστές ουρών αρκεί να γνωρίζουν τη θέση του πλησιέστερου δρομολογητή.

Γενική οργάνωση ενός συστήματος ουρών μηνυμάτων με δρομολογητές





Εφαρμογές των Συστημάτων Ουρών Μηνυμάτων

Τα συστήματα ουρών μηνυμάτων υποστηρίζουν εγγυημένη παράδοση μηνυμάτων, προτεραιότητες μηνυμάτων, αποδοτική πολυεκπομπή, εξισορρόπηση φορτίου, ανοχή σε βλάβες, κ.λπ.

General-purpose message queuing systems support a wide range of applications, including:

- Electronic mail.
- Workflow.
- Groupware.
- Batch Processing.

Most important message queuing application area:

Η ενοποίηση μιας συλλογής εφαρμογών βάσεων δεδομένων σε ένα πληροφοριακό σύστημα πολυβάσης δεδομένων



Διαπραγματευτές μηνυμάτων (message brokers)

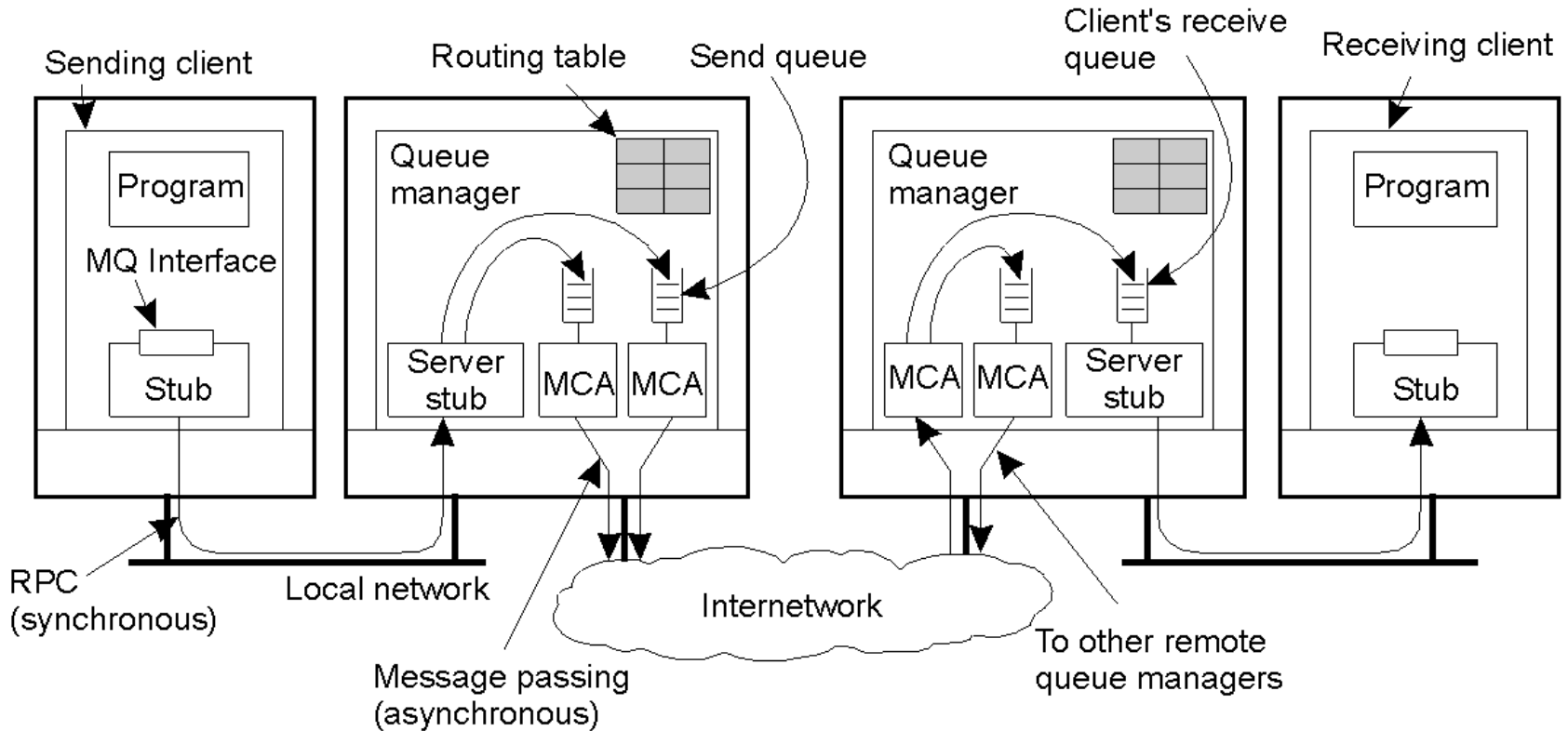
Συχνή η ανάγκη για ενοποίηση υφισταμένων και νέων εφαρμογών σε ένα **ενιαίο και συνεπές Distributed Information System (DIS)**.

Πρόβλημα: Διαφορετικές μορφές μηνυμάτων σε διαφορετικές εφαρμογές με αποτέλεσμα να ανταλλάσσονται μηνύματα σε μη κατανοητή μορφή.

Λύση: Χρήση **διαπραγματευτών μηνυμάτων** ήτοι εφαρμογές που συμπεριφέρονται ως πύλες του επιπέδου εφαρμογής, με κύριο σκοπό τη μετατροπή των εισερχόμενων μηνυμάτων σε μια μορφή που να μπορεί να γίνει κατανοητή από την εφαρμογή προορισμού.

Καρδιά των **διαπραγματευτών μηνυμάτων** μια βάση δεδομένων **κανόνων μετατροπής μηνυμάτων**.

Example: IBM MQSeries



- Η διαχείριση ουρών γίνεται από διαχειριστές ουρών (queue managers).
- Οι διαχειριστές συνδέονται ανά ζεύγη μέσω καναλιών μηνυμάτων.
- Ο χειρισμός καθενός από τα 2 άκρα ενός καναλιού μηνυμάτων γίνεται από ένα πράκτορα καναλιού μηνυμάτων (Message Channel Agent-MCA).

IBM MQSeries



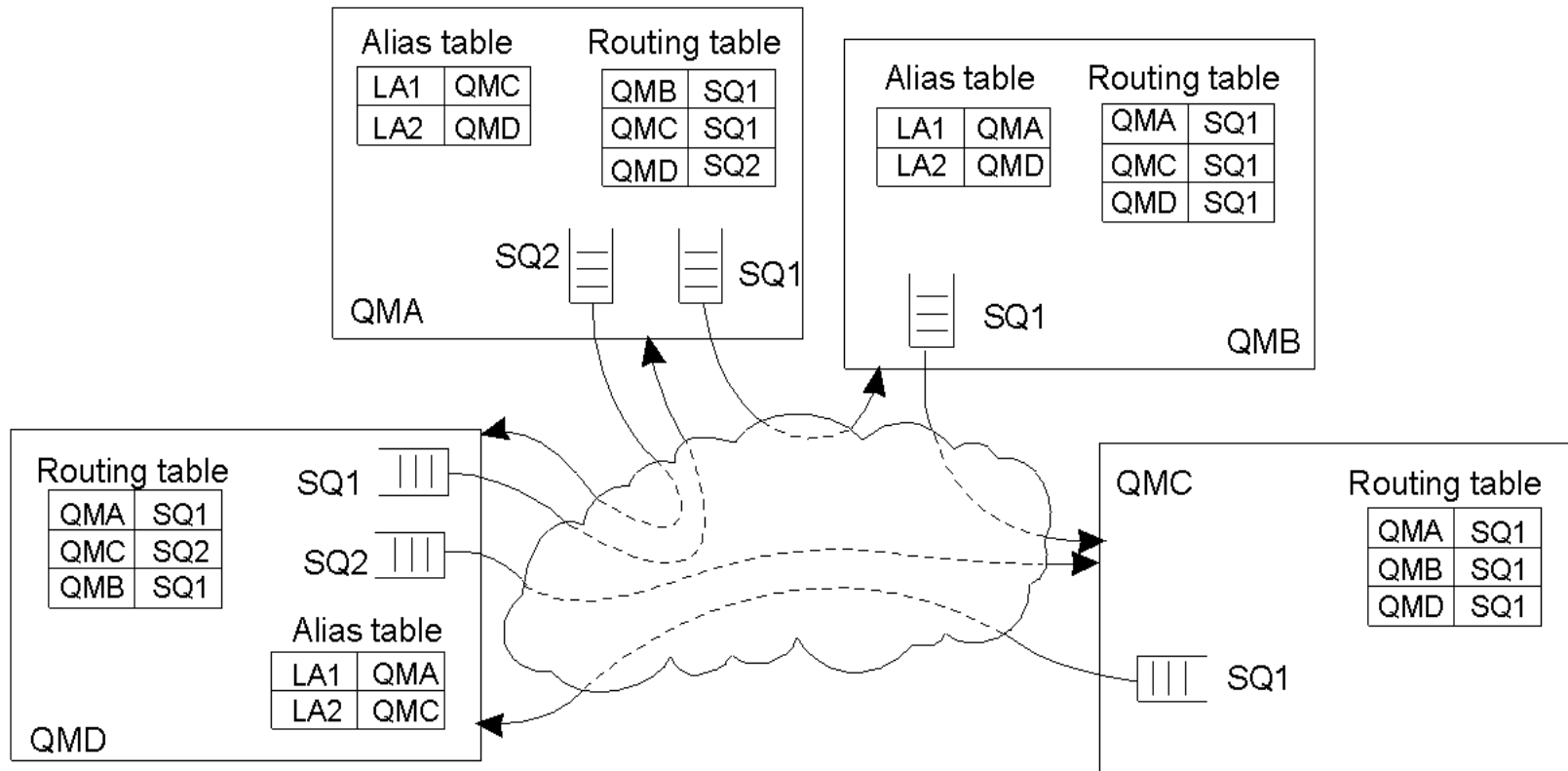
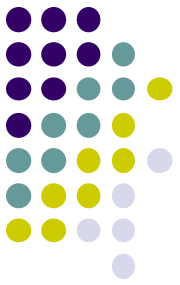
- Κάθε κανάλι μηνυμάτων αντιστοιχεί σε μία ουρά αποστολής και μία ουρά παραλαβής.
- Ένας MCA αποστολής ελέγχει την ουρά αποστολής, συσκευάζει μηνύματα σε πακέτα του επιπέδου μεταφοράς και το στέλνει στον αντίστοιχο MCA παραλαβής.
- Ένας MCA παραλαβής περιμένει για εισερχόμενο μήνυμα, το αποσυσκευάζει και το αποθηκεύει στην ουρά.
- Στην περίπτωση που η εκτέλεση των διαχειριστών ουρών και των εφαρμογών γίνεται σε διαφορετικά μηχανήματα, παρέχεται στην εφαρμογή διασύνδεση που υλοποιείται ως διαμεσολαβητής (proxy) που επικοινωνεί με το διαχειριστή με RPC.

Διευθυνσιοδότηση



- Οι διευθύνσεις αποτελούνται από το **όνομα του διαχειριστή ουρών** στον οποίο θα παραδοθεί το μήνυμα και το **όνομα της ουράς προορισμού** που υπάγεται σ' αυτόν τον διαχειριστή, όπου θα προσαρτηθεί το μήνυμα.
- Routing can be done using routing tables.
- Ένα ψευδώνυμο επιτρέπει να χρησιμοποιείται το ίδιο λογικό όνομα για μια ουρά ακόμη και αν αλλάξει ο διαχειριστής ουρών που τη χειρίζεται.
- Για να αλλάξουμε το όνομα ενός διαχειριστή ουρών απαιτείται να ενημερώσουμε τον πίνακα ψευδωνύμων σε όλους τους διαχειριστές. Οι εφαρμογές μένουν ανεπηρέαστες.

Message Queue Interface (MQI)



Παράδειγμα: μια εφαρμογή συνδεδεμένη στον διαχειριστή ουρών QMA αναφέρεται σε έναν απομακρυσμένο διαχειριστή ουρών χρησιμοποιώντας το ψευδώνυμο LA1.

Ρευματοστρεφής επικοινωνία (Stream-Oriented Communications)



RPC, RMI, MOM: ο χρόνος δεν έχει καμιά επίπτωση στην ορθότητα της επικοινωνίας

Όμως, τα συνεχή μέσα (ήχος, βίντεο, κινούμενα γραφικά) αποτελούν **εξαρτώμενα από το χρόνο ρεύματα δεδομένων (*time-dependent data streams*)** δηλαδή η χρονική στιγμή κατά την οποία πραγματοποιείται η επικοινωνία έχει άμεση επίπτωση στην ορθότητα ολοκλήρωσής της.

Απαιτούνται ειδικά πρωτόκολλα για την υποστήριξη της μετάδοσης των συνεχών μέσων (**ρευματοστρεφής επικοινωνία**)

Example: video: 30 frames per second (30-40 msec per image).

Ρεύμα δεδομένων



Απλό Ρεύμα - Simple Stream – one single sequence of data, for example: voice.

Σύνθετο Ρεύμα - Complex Stream – several sequences of data (υπορεύματα -substreams) that are “related” by time.

E.g., a lipsynchronized movie, with sound and pictures, together with sub-titles ...

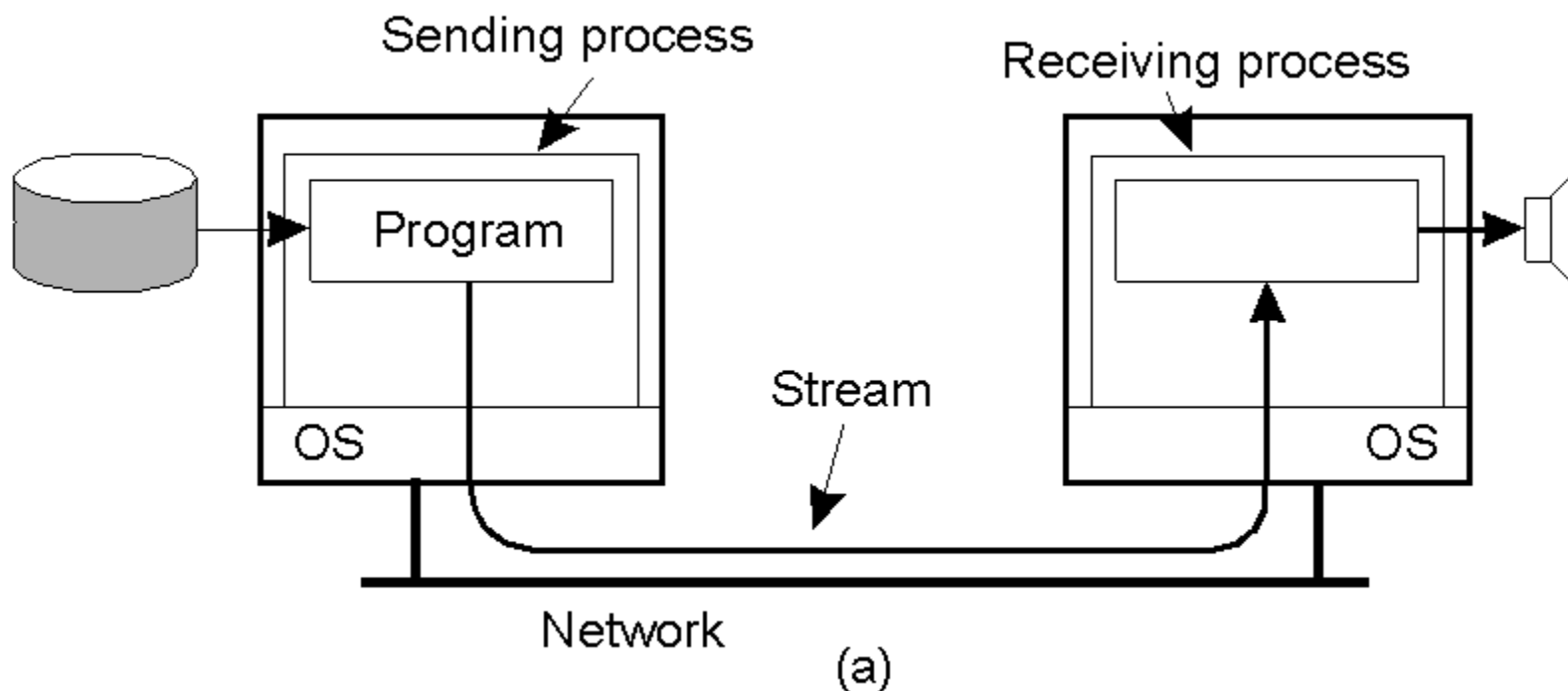
This leads to data synchronization problems ... which are not at all easy to deal with.

Ρεύμα δεδομένων



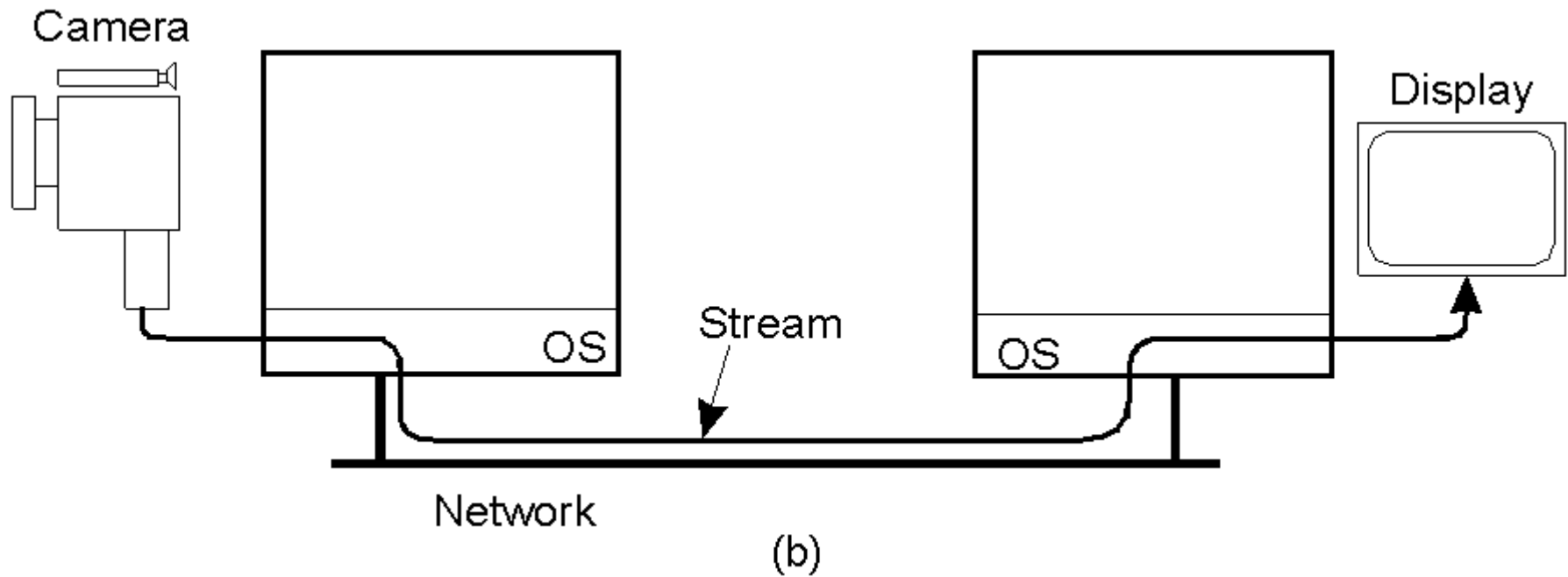
Μπορεί να θεωρηθεί ως μια εικονική σύνδεση μεταξύ μιας πηγής και ενός προορισμού (διεργασία ή συσκευή).

Κατανεμημένα συστήματα πολυμέσων: παρέχουν υποστήριξη για την εγκατάσταση απευθείας σύνδεσης μεταξύ πηγής και προορισμού



Ρεύμα δεδομένων

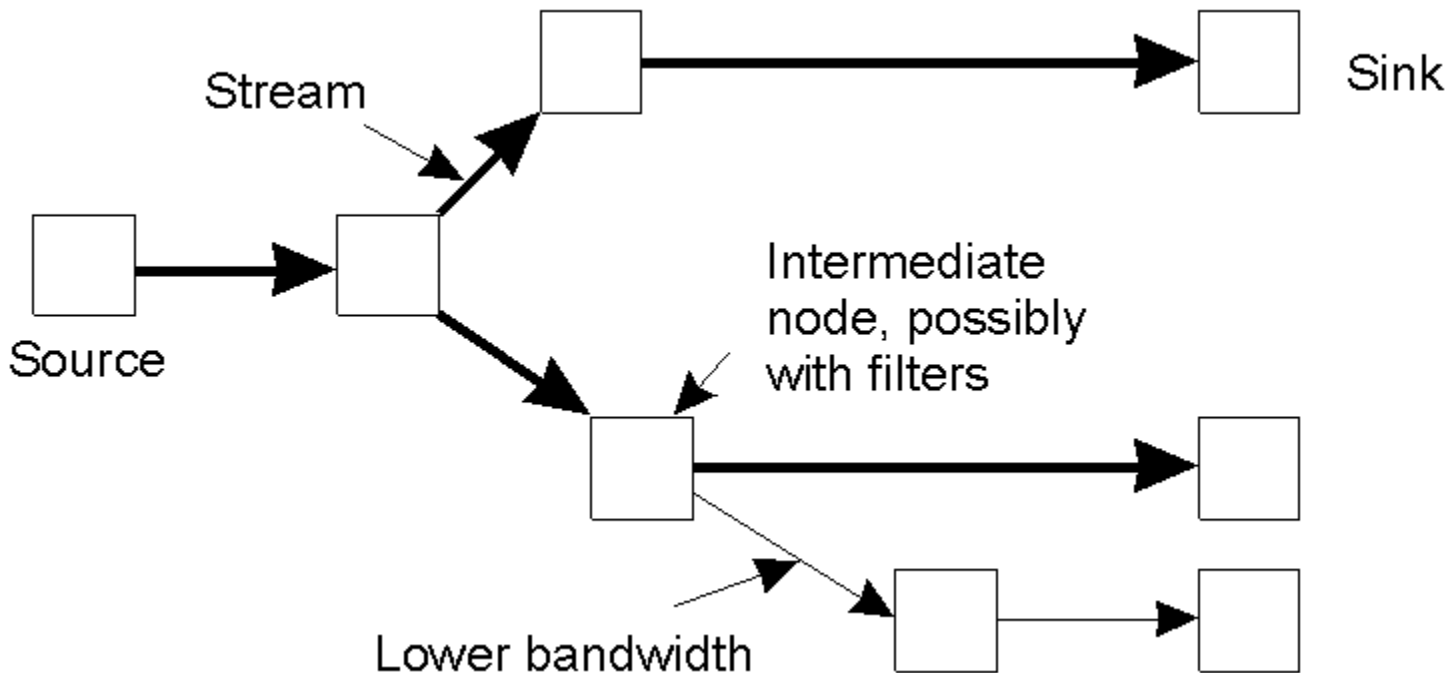
Ρεύμα video μπορεί να προωθείται απευθείας σε μια συσκευή απεικόνισης



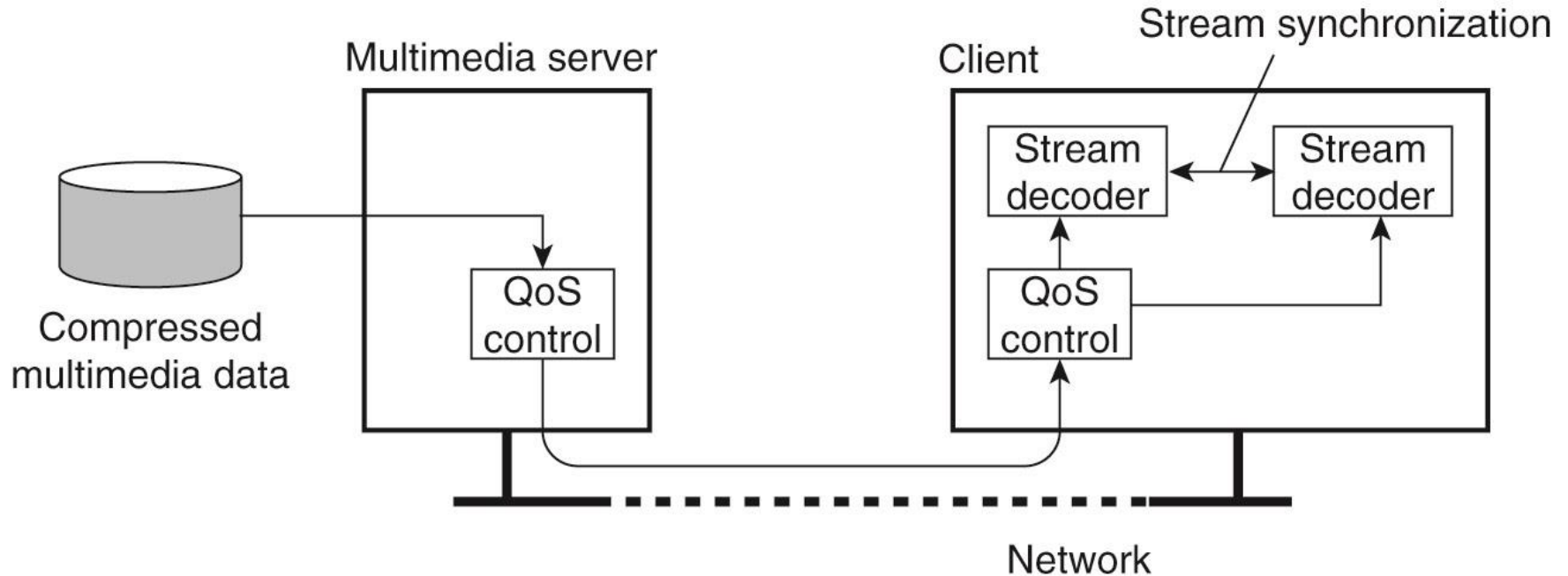
Πολυεκπομπή ρεύματος



Χρήση φίλτρων για την προσαρμογή της ποιότητας του εισερχόμενου ρεύματος διαφορετικά για τα εξερχόμενα ρεύματα.



Ρεύμα δεδομένων



A general architecture for streaming stored multimedia data over a network.

Καταστάσεις μετάδοσης (transmission modes)



Κατάσταση ασύγχρονης μετάδοσης - Asynchronous transmission mode – δεν υπάρχουν περιορισμοί χρονισμού για το πότε θα πρέπει να γίνεται η μετάδοση (e.g., File Transfer).

Κατάσταση σύγχρονης μετάδοσης - Synchronous transmission mode – ορίζεται **μέγιστη καθυστέρηση μεταφοράς από άκρο σε άκρο (maximum end-to-end delay)** για κάθε μονάδα ενός ρεύματος δεδομένων αλλά μπορεί να γίνει γρηγορότερη μετάδοση

Κατάσταση ισόχρονης μετάδοσης - Isochronous transmission mode – data transferred “on time” – ορίζεται μέγιστη και ελάχιστη καθυστέρηση μεταφοράς από άκρο σε άκρο που ονομάζεται **οριοθετημένη διακύμανση καθυστέρησης (“bounded jitter”)**.

Εφαρμογές και οι απαιτήσεις τους



- **Streaming** (μετάδοση βίντεο σε πραγματικό χρόνο και αναπαραγωγή χωρίς αποθήκευση)
 - **Εύρος ζώνης (bandwidth)**: Ίσο με το bitrate του βίντεο μετά από συμπίεση.
 - **Αποθήκευση (storage)**: Δεν χρειάζεται αποθήκευση
 - **Καθυστέρηση για αναπαραγωγή**: Μικρή (έως ότου συμπληρωθεί ένα αυτόνομο τμήμα για αποκωδικοποίηση)
 - **Ανάλυση**: Εξαρτάται από το διαθέσιμο εύρος ζώνης
- **Downloading** (Video on Demand):
 - **Εύρος ζώνης (bandwidth)**: Μπορεί να είναι πάρα πολύ μικρό (από dial-up σύνδεση 56kbps έως LAN 100 Mbps).
 - **Αποθήκευση (storage)**: Τόσο όσο να μπορεί να αποθηκευτεί ολόκληρο το αρχείο βίντεο.
 - **Καθυστέρηση για αναπαραγωγή**: Ίσο με το χρόνο που χρειάζεται για να γίνει download ολόκληρο το αρχείο
 - **Ανάλυση**: Εξαρτάται από το διαθέσιμο αποθηκευτικό χώρο

Ρεύματα Δεδομένων και Ποιότητα Παροχής Υπηρεσιών (QoS)



Η ποιότητα παροχής υπηρεσιών εκφράζεται με προδιαγραφές ροής για το ρεύμα δεδομένων:

- **καθορισμός χρονικών σχέσεων μεταξύ των μονάδων δεδομένων**
- **απαιτήσεις εξυπηρέτησης όπως:**
 - το απαιτούμενο εύρος ζώνης (π.χ. για ροή βίντεο 1,5 Mbps)
 - ο μέγιστος ρυθμός απωλειών (π.χ. 1 καρτέ ανά 100)
 - η μέγιστη καθυστέρηση
 - η μέγιστη διακύμανση καθυστέρησης
 - ποιότητα εγγύησης (ένας αριθμός που προσδιορίζει πόσο σοβαρά πρέπει να ληφθούν από το σύστημα επικοινωνίας οι απαιτήσεις εξυπηρέτησης ώστε να αποφασισθεί αν θα εγκατασταθεί ή όχι το ρεύμα)

Εγκατάσταση Ρεύματος



Αφού μέσω των προδιαγραφών ροής περιγραφεί το ρεύμα δεδομένων, **το κατανεμημένο σύστημα εκχωρεί πόρους:**

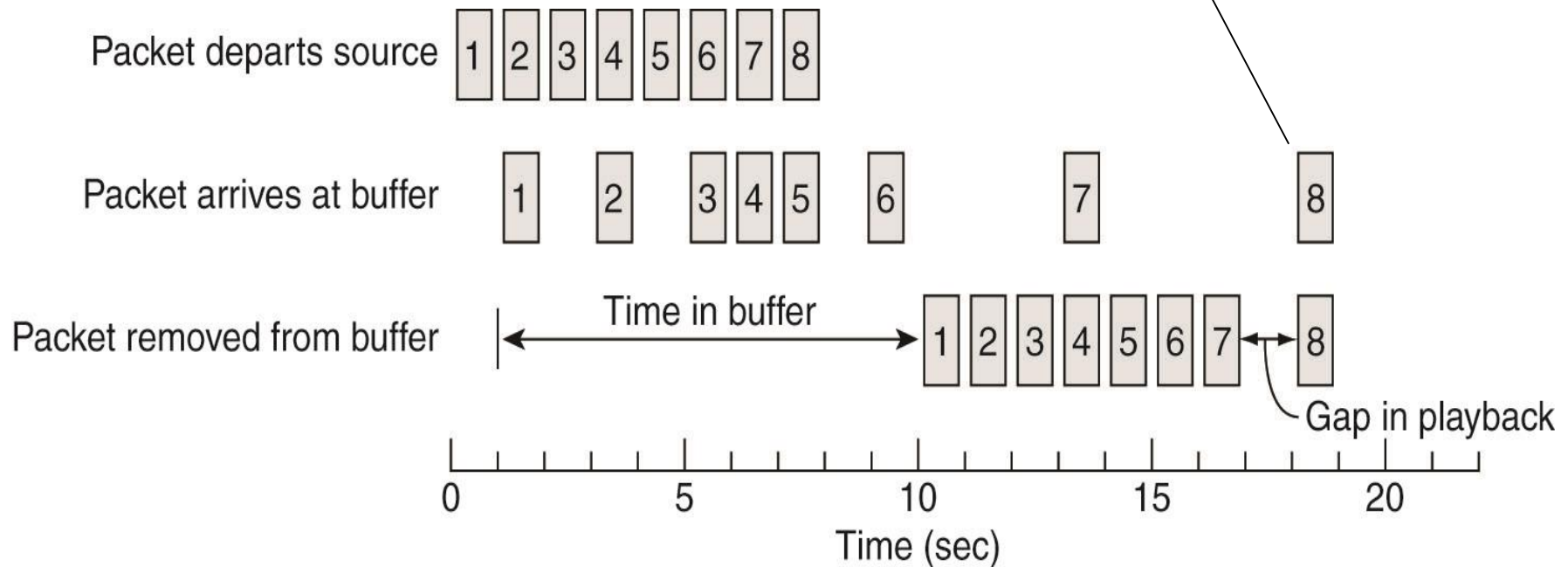
- **δεύσμευση** εύρους ζώνης μέσω κατάλληλου χρονοπρογραμματισμού που εξασφαλίζει τη μετάδοση των μονάδων δεδομένων,
- **εκχώριση προσωρινής μνήμης** στους δρομολογητές και τον παραλήπτη,
- **επεξεργαστική ικανότητα** μέσω κατάλληλου προγραμματισμού της χρήσης των CPU

Όστε να είναι δυνατή η εγκατάσταση του ρεύματος δεδομένων που ικανοποιεί τις δεδομένες απαιτήσεις QoS.



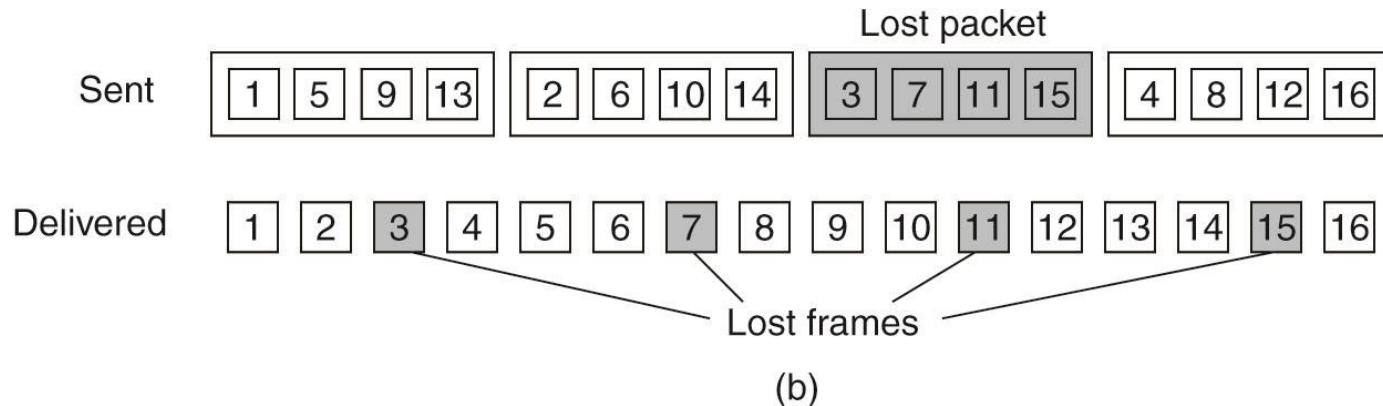
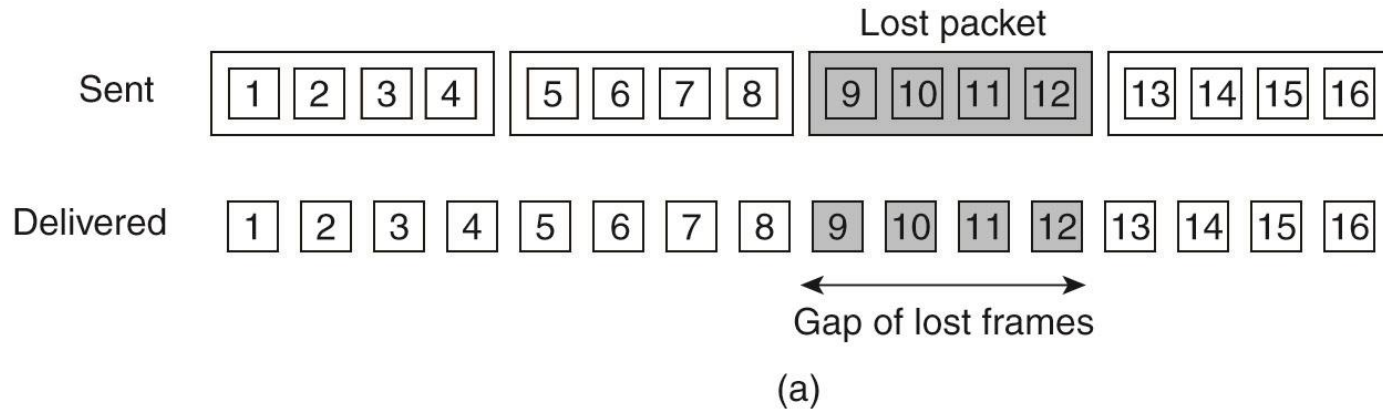
Εξασφάλιση QoS- Audio&Video on Demand

Η διακύμανση στην καθυστέρηση είναι μεγαλύτερη απ' ότι αναμενόταν με αποτέλεσμα το πακέτο να φτάσει καθυστερημένο



- Η διακύμανση στην καθυστέρηση έχει ως αποτέλεσμα τα πακέτα να φτάνουν καθυστερημένα.
- Η αποθήκευση στην ενδιάμεση μνήμη χρησιμοποιείται για τη βελτίωση της ποιότητας της υπηρεσίας.

Εξασφάλιση QoS



Απώλεια πακέτων στην (a) non interleaved transmission και (b) interleaved transmission.

Resource reSerVation Protocol (RSVP)



Το RSVP είναι πρωτόκολλο ελέγχου του επιπέδου μεταφοράς που χρησιμοποιείται για τη δέσμευση πόρων σε δρομολογητές δικτύου, επιτρέποντας την παροχή ποιότητας υπηρεσίας (QoS) για εφαρμογές όπως video streaming, VoIP και λοιπές απαιτητικές υπηρεσίες δικτύου.

1. Δημιουργία Αιτήματος Δέσμευσης Πόρων

- Ο **αποστολέας** (π.χ., ένας video streaming server) στέλνει ένα **PATN μήνυμα** προς τον παραλήπτη.
- Αυτό το μήνυμα περιέχει τις **προδιαγραφές ροής** του ρεύματος δεδομένων, όπως το απαιτούμενο εύρος ζώνης, η καθυστέρηση κ.λπ.
- Το PATN μήνυμα περνά από τους ενδιαμέσους δρομολογητές, οι οποίοι ενημερώνουν την κατάσταση τους αλλά **δεν δεσμεύουν ακόμα πόρους**.

2. Ενεργοποίηση RSVP από τον Αποδέκτη

- Ο **παραλήπτης** (π.χ., ένας χρήστης που παρακολουθεί το βίντεο) λαμβάνει το PATN μήνυμα και στέλνει πίσω ένα **RESV μήνυμα** προς τον αποστολέα.
- Το RESV μήνυμα ζητά **επίσημα τη δέσμευση πόρων**, ακολουθώντας την ίδια διαδρομή αντίστροφα.

Resource reSerVation Protocol (RSVP)



3. Έλεγχος και Δέσμευση Πόρων

Κάθε ενδιαμέσος **δρομολογητής** που λαμβάνει το RSVP μήνυμα εκτελεί 2 ελέγχους:

1. **Έλεγχος Αποδοχής (Admission Control)**: Ελέγχει αν υπάρχουν **διαθέσιμοι πόροι** (π.χ., εύρος ζώνης) για να ικανοποιήσει το αίτημα.
2. **Έλεγχος Πολιτικής (Policy Control)**: Ελέγχει αν ο **παραλήπτης έχει άδεια** για τη δέσμευση των πόρων (π.χ., βάσει κανόνων δικτύου ή προτεραιοτήτων χρήσης).

Αν και οι δύο έλεγχοι είναι επιτυχείς, **οι πόροι δεσμεύονται** κατά μήκος της διαδρομής. Αν κάποιος δρομολογητής δεν έχει αρκετούς πόρους, η αίτηση μπορεί να απορριφθεί ή να τροποποιηθεί.

4. Μετάφραση Παραμέτρων QoS στο Επίπεδο Συνδέσμου

- Η διεργασία RSVP **μεταφράζει τις παραμέτρους QoS** σε ρυθμίσεις που έχουν νόημα στο **επίπεδο συνδέσμου δεδομένων** (π.χ. τιμές παραμέτρων του ATM).

5. Ανανεώσεις & Απελευθέρωση Πόρων

- Οι δεσμεύσεις πόρων **ανανεώνονται περιοδικά**. Αν σταματήσει η χρήση της 4/4/2025 υπηρεσίας, οι δεσμευμένοι πόροι **απελευθερώνονται αυτόματα**. 50

Resource reSerVation Protocol (RSVP)



Υπομονάδα Ελέγχου Αποδοχής (Admission Control)

- Όταν ένα μήνυμα **RESV** φτάσει σε έναν δρομολογητή, η διεργασία RSVP το μεταβιβάζει στην υπομονάδα ελέγχου αποδοχής η οποία ελέγχει **αν υπάρχουν επαρκείς πόροι** (π.χ., εύρος ζώνης) για να ικανοποιηθεί το αίτημα του αποδέκτη.
- Αν οι διαθέσιμοι πόροι είναι ανεπαρκείς, το αίτημα απορρίπτεται.

Υπομονάδα Ελέγχου Πολιτικής (Policy Control)

- Η αίτηση ελέγχεται και η μονάδα επαληθεύει **αν ο αποδέκτης έχει την απαιτούμενη άδεια** για δέσμευση των ζητούμενων πόρων.
- Αυτό είναι σημαντικό σε δίκτυα όπου υπάρχουν περιορισμοί ή προτεραιότητες χρήσης (π.χ., VIP χρήστες, εταιρικές πολιτικές).

Δέσμευση Πόρων

- Αν και οι δύο έλεγχοι ολοκληρωθούν με επιτυχία, ο δρομολογητής **δεσμεύει τους απαιτούμενους πόρους**. Η διαδικασία αυτή συνεχίζεται για όλους τους ενδιαμέσους δρομολογητές, μέχρι να διαμορφωθεί η πλήρης διαδρομή.

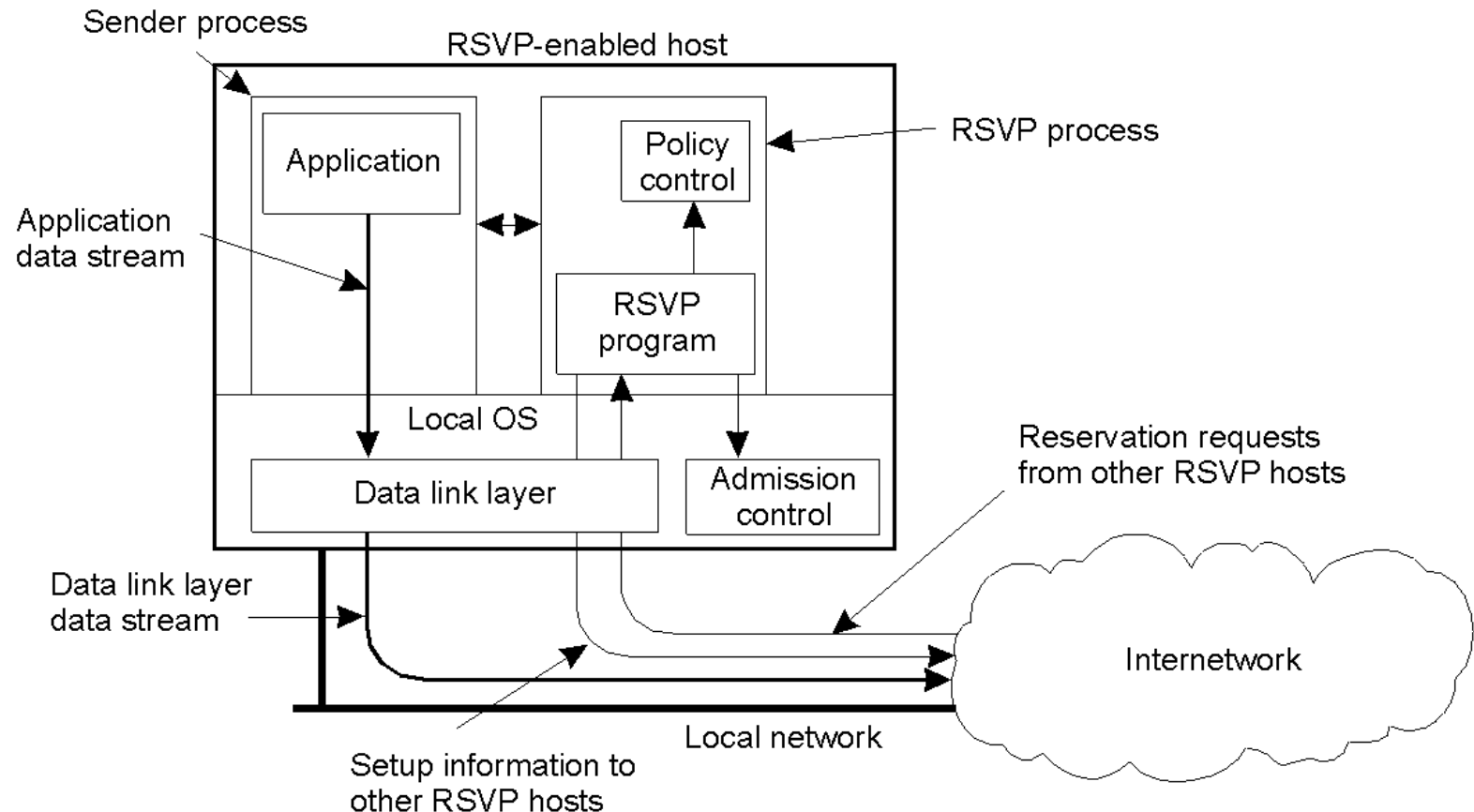
Μετάφραση Παραμέτρων QoS στο Επίπεδο Συνδέσμου Δεδομένων

Η διεργασία RSVP μεταφράζει τις παραμέτρους **QoS** σε αντίστοιχες ρυθμίσεις για το υποκείμενο επίπεδο σύνδεσης δεδομένων (π.χ. τιμές παραμέτρων του ATM).⁵¹



Resource reSerVation Protocol (RSVP)

The basic organization of RSVP for resource reservation in a distributed system.



Συγχρονισμός ρευμάτων



Διατήρηση χρονικών σχέσεων μεταξύ τους

- Υποστηρίζεται από LDU: λογικές μονάδες δεδομένων ενός μέσου

Στα συνεχή ρεύματα έχουμε χρονικές συσχετίσεις ανάμεσα στις LDU του ίδιου μέσου

Στα διακριτά μέσα δεν έχουμε χρονικές συσχετίσεις ανάμεσα στις LDU του ίδιου μέσου

- Τα LDU ενός συνεχούς ρεύματος μπορούν να ορισθούν σε διάφορα επίπεδα. Π.χ. στο video LDU μπορεί να είναι τα GOP, τα καρέ, τα τμήματα, τα μακρομπλοκ ή τα μπλοκ.



Παράδειγμα: Ρεύμα μονοκάναλου ήχου ποιότητας CD εμφανίζεται ως ακολουθία δειγμάτων των 16 bits. Με συχνότητα δειγματοληψίας 44100 Hz (δείγματα ανά δευτερόλεπτο), ο συγχρονισμός με άλλα ρεύματα θα μπορούσε να γίνεται κάθε 23 μsec .

Ρεύμα βίντεο εικόνων 320 x 240 pixels, το καθένα κωδικοποιημένο με 1 byte, οδηγεί σε μονάδες δεδομένων βίντεο των 76800 bytes η κάθε μία. Οι εικόνες εμφανίζονται με συχνότητα 30 Hz ήτοι μία κάθε 33 msec.

Για να επιτευχθεί συγχρονισμός βίντεο και ήχου θα πρέπει τα ηχητικά δείγματα να είναι σε μονάδες των 2870 bytes (1435 δείγματα των 16 bits) οι οποίες έχουν διάρκεια 33msec.

Ο συγχρονισμός μπορεί να επιτευχθεί με απλή εναλλαγή μεταξύ της ανάγνωσης μιας εικόνας και μιας μονάδας ήχου (1435 δείγματα) κάθε 33 msec.



Είδη Συγχρονισμού

- **Συγχρονισμός ενός αντικειμένου**
 - Αφορά τη χρονική ακρίβεια της παρουσίασης των διαδοχικών LDU
 - Απαίτηση είναι η αποφυγή διαταραχής στο χρόνο παρουσίασης
 - Χρονικές σχέσεις ανάμεσα στα LDU ενός συνεχούς μέσου (π.χ. τα καρέ ενός βίντεο με ρυθμό αναπαραγωγής 30Hz πρέπει να παρουσιάζονται κάθε 33msec)
- **Συγχρονισμός πολλαπλών αντικειμένων**
 - Χρονικές σχέσεις ανάμεσα σε πολλαπλά μέσα
 - Συγχρονισμός μετά από διαφορετικές καθυστερήσεις

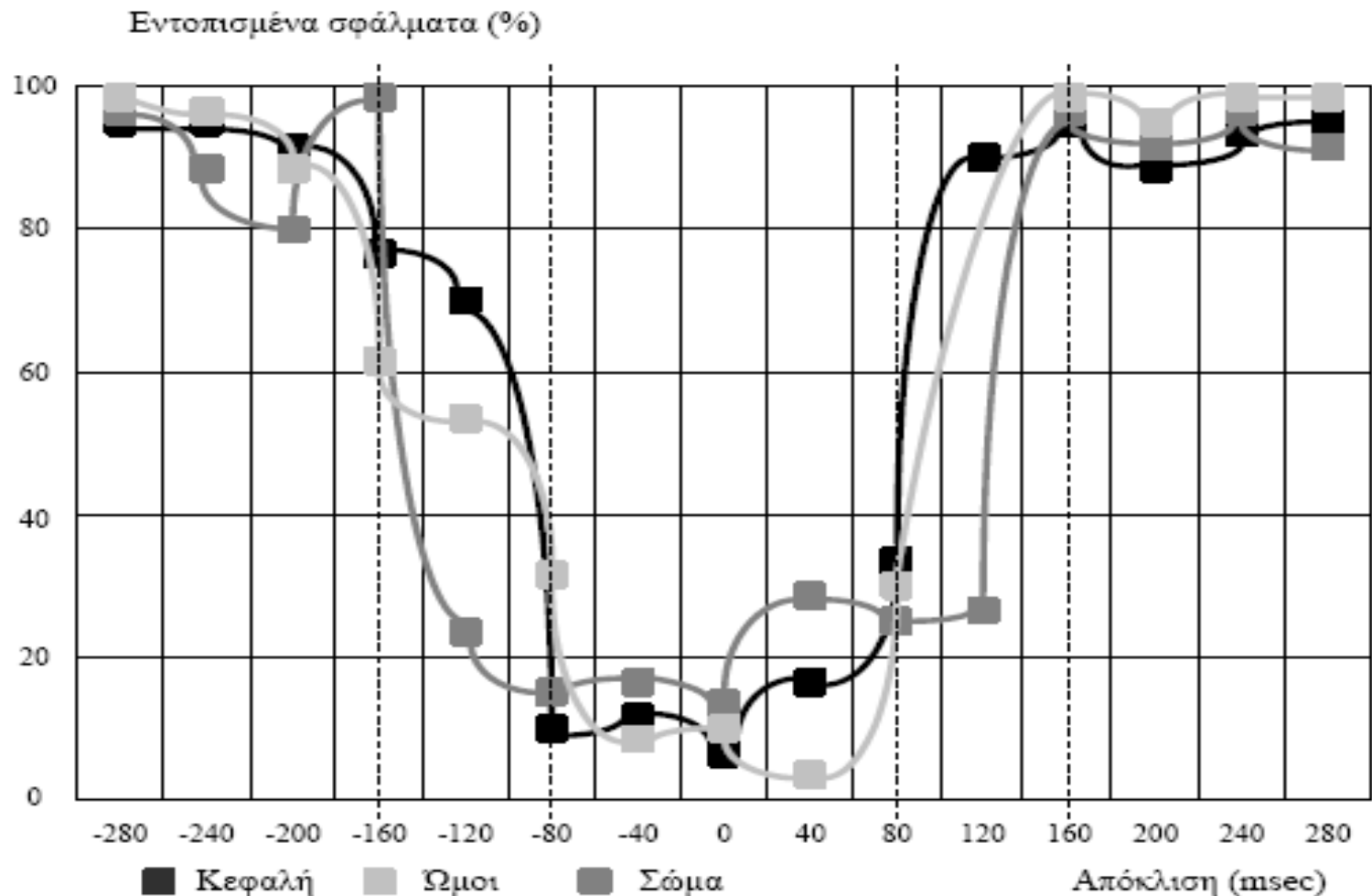
Απαιτήσεις συγχρονισμού



- **Συγχρονισμός χειλιών**
 - Βίντεο (χείλη) και ήχος (φωνή)
 - Απόκλιση είναι η χρονική διαφορά στην παρουσίαση των LDU ήχου και βίντεο που κανονικά πρέπει να παρουσιάζονται ταυτόχρονα
 - Θετική απόκλιση: βίντεο έπεται του ήχου
 - Αρνητική απόκλιση: ήχος έπεται του βίντεο
- **Εμπειρική μελέτη**
 - Παρουσιαστής ειδήσεων, όψη κεφαλής, ώμων, σώματος
 - Παρόμοια αποτελέσματα (καμπύλες)
 - Ελαφρά μεγαλύτερη ενόχληση σε όψη κεφαλής

Απαιτήσεις συγχρονισμού

- ανοχή σφαλμάτων συγχρονισμού χειλιών





Απαιτήσεις συγχρονισμού

- **Ανοχή στην απόκλιση**
 - Πιο ενοχλητική η θετική απόκλιση
 - Πιο απότομη αύξηση καμπυλών στα δεξιά
 - Είμαστε πιο ανεκτικοί στην αρνητική απόκλιση διότι η ταχύτητα του φωτός είναι μεγαλύτερη της ταχύτητας του ήχου
 - Απόσταση 20 m, αρνητική απόκλιση 60 ms
 - Η αρνητική απόκλιση είναι καθημερινό φαινόμενο
- **Εντός συγχρονισμού: -80 ms έως $+80$ ms**
 - Δεν έγιναν αντιληπτές ή κρίθηκαν αποδεκτές
- **Εκτός συγχρονισμού: πάνω από ± 160 ms**
 - Έγιναν αντιληπτές και κρίθηκαν ενοχλητικές.
- **Μεταβατικές: ± 80 έως 160 ms**
 - Αποδοχή ανάλογα με παρατηρούμενη εικόνα

Συγχρονισμός



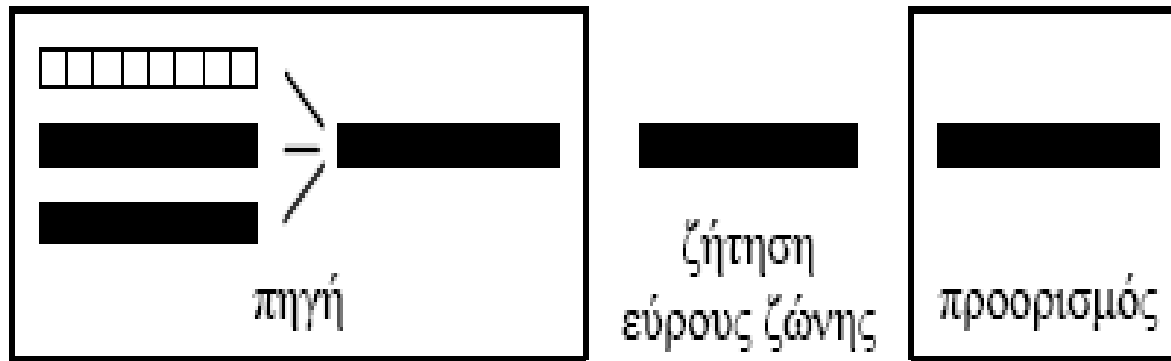
Κατανεμημένο σύστημα πολυμέσων

- Αυξημένη καθυστέρηση και διαταραχή λόγω μετάδοσης
- Πρέπει εκτός από τα μέσα να μεταδίδονται και οι προδιαγραφές συγχρονισμού τους.

Πότε μεταδίδονται οι προδιαγραφές;

- Συγχρονισμός στην πηγή
- Συγχρονισμός στον προορισμό

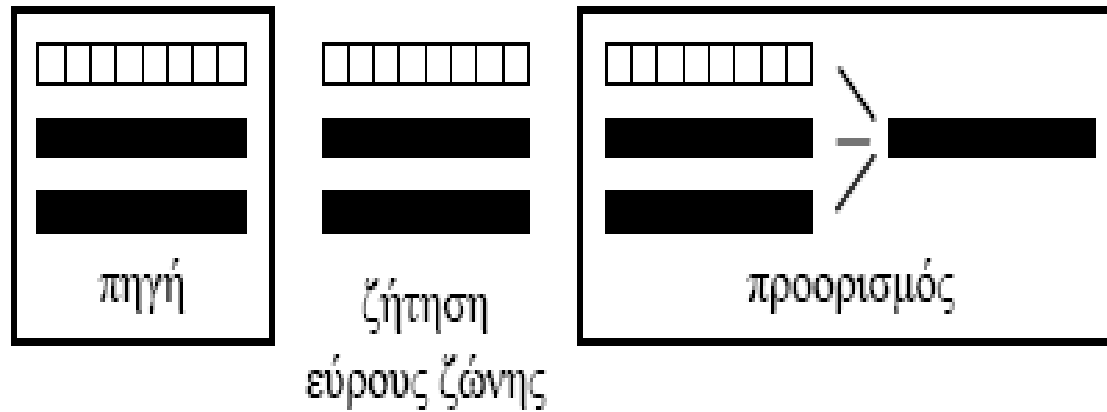
Συγχρονισμός



Συγχρονισμός στην πηγή

- Αρκεί ένα κανάλι για όλες τις πληροφορίες
- Οι απαιτήσεις συγχρονισμού ορίζονται από την πηγή κατά την πολύπλεξη χωρίς δυνατότητα παρέμβασης του παραλήπτη.

Συγχρονισμός

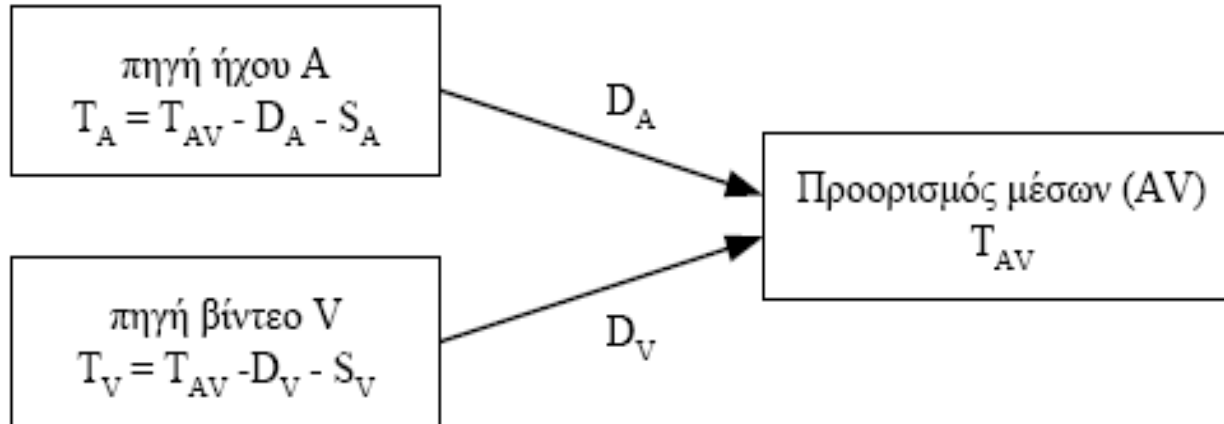


Συγχρονισμός στον προορισμό

- Συγχρονισμός πριν την παρουσίαση
- Δυνατότητα παρέμβασης στον προορισμό

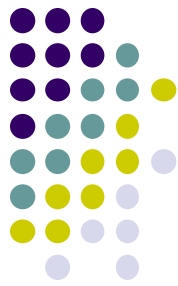


Συγχρονισμός



- Ακρίβεια συγχρονισμού ρολογιών
- Παρουσίαση ήχου και βίντεο τη στιγμή T_{AV}
 - Ήχος: αρχίζει τη στιγμή $T_A = T_{AV} - D_A - S_A$
 - Βίντεο: αρχίζει τη στιγμή $T_V = T_{AV} - D_V - S_V$
 - D_A, D_V : καθυστέρηση μετάδοσης
 - S_A, S_V : απόκλιση ανάμεσα στα ρολόγια
- Εκτίμηση άνω ορίου απόκλισης
 - Εκχώρηση ενταμιευτών και έγκαιρη εκκίνηση
 - Έγκαιρη εκκίνηση μετάδοσης

Μηχανισμός συγχρονισμού – Υποστηρίζεται από διασυνδέσεις του ενδιάμεσου λογισμικού



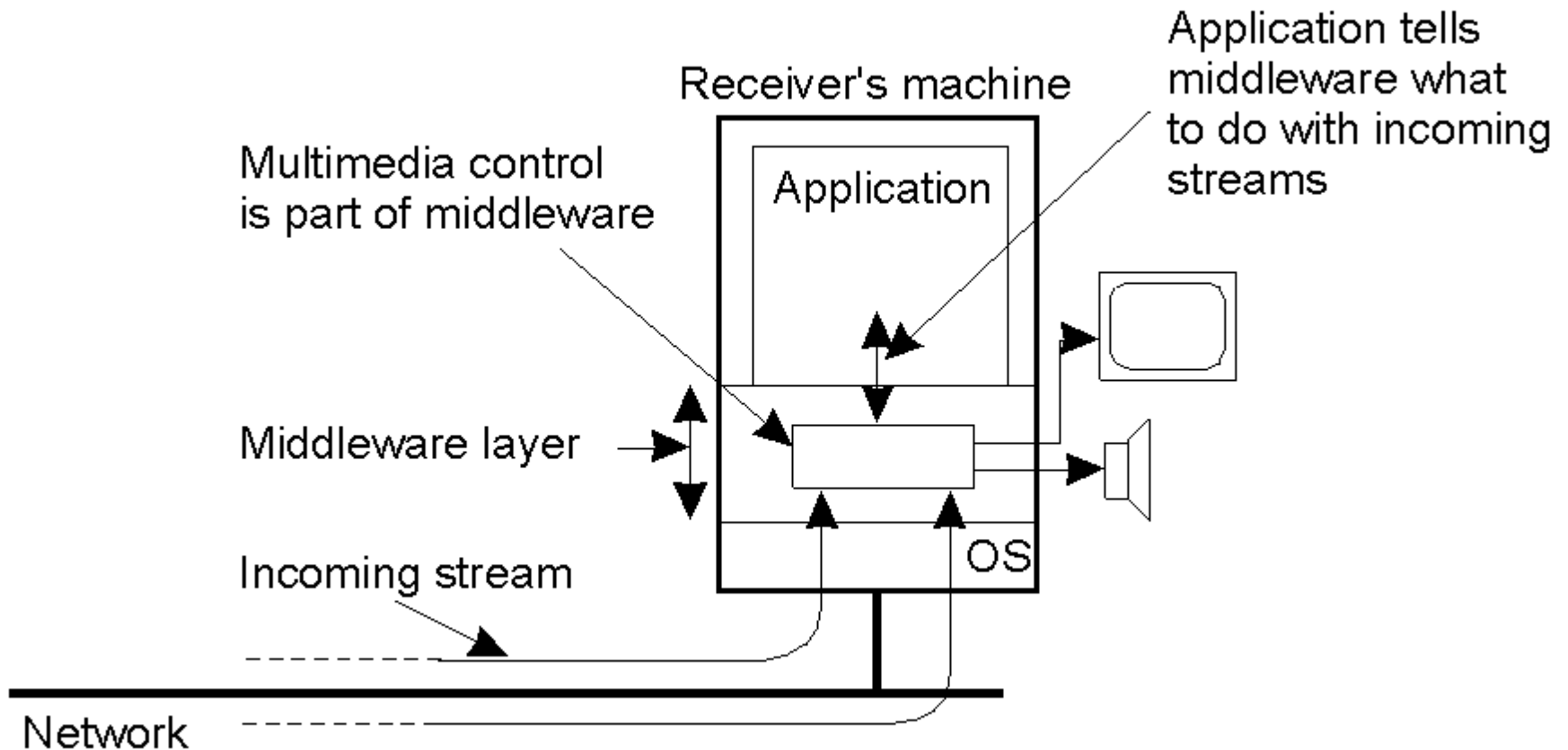
Το ενδιάμεσο λογισμικό πολυμέσων παρέχει διασυνδέσεις για τον έλεγχο των ρευμάτων ήχου και βίντεο και των συσκευών εξόδου.

Με αυτές τις διασυνδέσεις ελέγχου, μπορεί να γραφεί απλό πρόγραμμα εποπτείας που θα αποτελείται από 2 χειριστές, έναν για κάθε ρεύμα.

Οι χειριστές ελέγχουν μαζί αν τα ρεύματα βίντεο και ήχου είναι αρκετά συγχρονισμένα και αν χρειάζεται, να προσαρμόζουν το ρυθμό με τον οποίο παρουσιάζονται οι μονάδες βίντεο και ήχου.



Μηχανισμός συγχρονισμού – Υποστηρίζεται από διασυνδέσεις του ενδιάμεσου λογισμικού





Tiger Video Server

- **Tiger Video Server:** Πλατφόρμα που χρησιμοποιείται για τη διαχείριση, επεξεργασία και μετάδοση βίντεο σε ζωντανό χρόνο.
- Χρησιμοποιείται για εφαρμογές που απαιτούν υψηλή ποιότητα μετάδοσης και διαχείρισης βίντεο, όπως τηλεοπτικοί σταθμοί, επιχειρήσεις με ανάγκες για ζωντανή ροή περιεχομένου.

Tiger Video Server

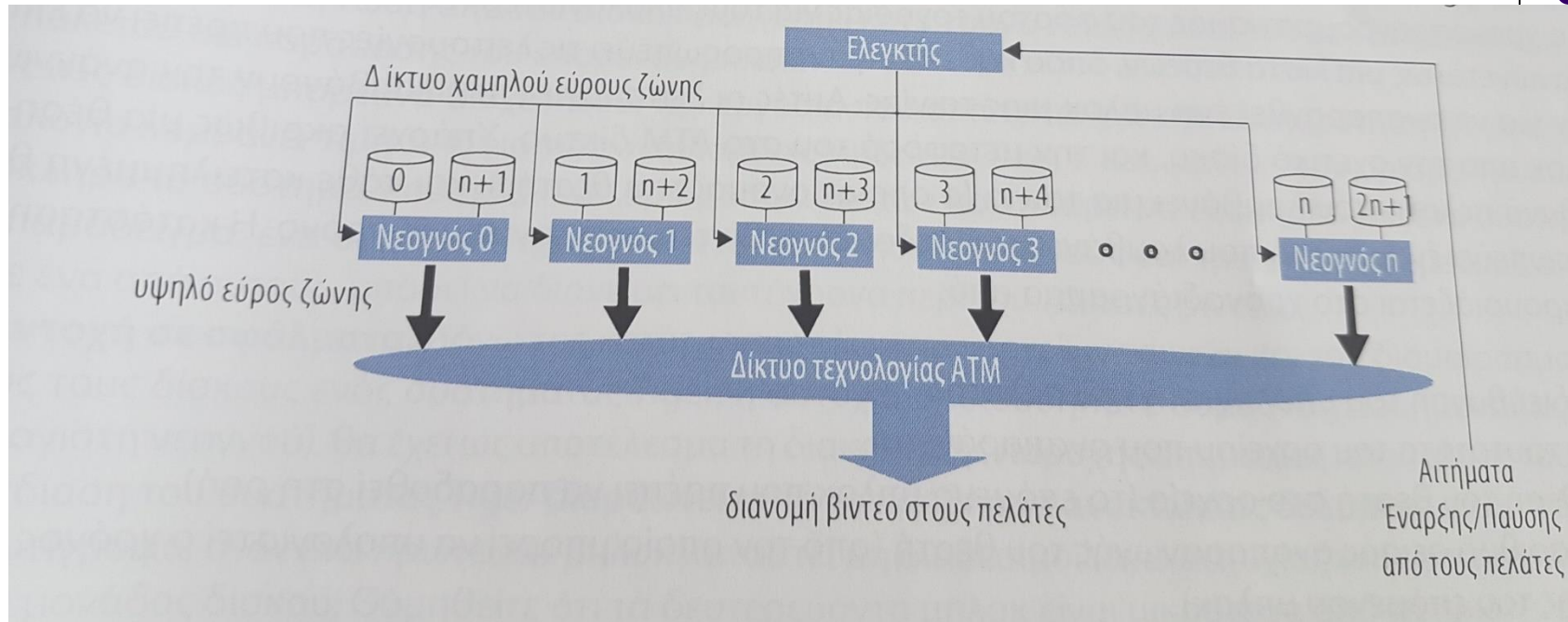
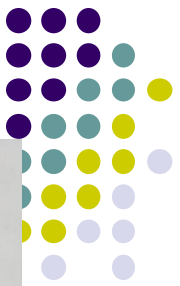


- Παρέχει ταυτόχρονα πολλαπλές ροές video σε πραγματικό χρόνο (π.χ. παροχή ταινιών σε συνδρομητικούς πελάτες)

Στόχοι:

- Η ροή θα πρέπει να αρχίζει μέσα σε λίγα δευτερόλεπτα από την έκδοση του αιτήματος
- Υποστήριξη ενεργειών όπως παύση και μετακίνηση σε προηγούμενη ή επόμενη σκηνή της ταινίας
- Ίδια ταινία μπορεί να είναι αντικείμενο ταυτόχρονης προβολής ως αποτέλεσμα ασύγχρονων αιτημάτων
- Επεκτάσιμη αρχιτεκτονική για υποστήριξη έως 10000 ταυτόχρονων χρηστών
- *Ποιότητα υπηρεσίας:* Οι ροές βίντεο να παρέχονται σε σταθερό ρυθμό με μέγιστη διακύμανση που εξαρτάται από την μικρή αποθηκευτική ικανότητα των χρηστών και χαμηλό ρυθμό απώλειας.

Αρχιτεκτονική του Tiger



Νεογνοί (Cubs): Υπολογιστές με τον ίδιο αριθμό μονάδων δίσκων

Ελεγκτής: Χειρίζεται αιτήματα πελατών και διαχειρίζεται το χρονοδιάγραμμα εργασιών των νεογνών

Κατάτμηση: Κάθε ταινία αποθηκεύεται σε κομμάτια video (μπλοκ) μεγέθους 0,5 MB (περίπου 7000) σε όλους τους δίσκους με τη σειρά των αριθμών του δίσκου, αναδιπλώνοντας μετά από $n+1$ κομμάτια

Αρχιτεκτονική του Tiger



Καθρεπτισμός: Κάθε κομμάτι video διαιρείται σε d δευτερεύοντα κομμάτια. Το d καλείται **συντελεστής αποσυμπλεγματοποίησης** και παίρνει τιμές από 4 έως 8.

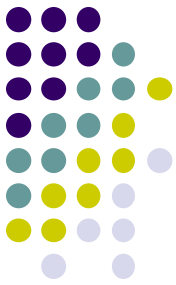
Τα δευτερεύοντα κομμάτια ενός κομματιού που είναι αποθηκευμένο στον δίσκο i αποθηκεύονται στους δίσκους $i+1$ έως $i+d$. Αν ο αριθμός των νεογνών είναι μεγαλύτερος από d , κανένας από αυτούς τους δίσκους δεν είναι στον ίδιο νεογνό με τον δίσκο i .

Αν κάποιος νεογνός αποτύχει, ο φόρτος για την παροχή δεδομένων από τα κομμάτια video που βρίσκονται σε αυτόν, μοιράζεται σε d νεογνούς.

Χρονοπρογραμματισμός στο Tiger

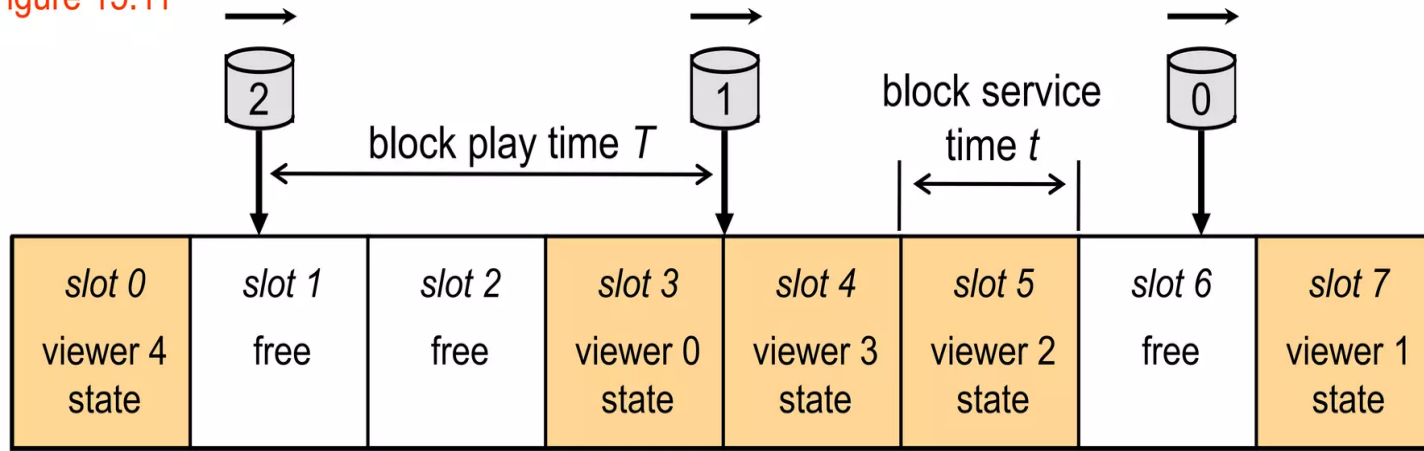


- Το χρονοδιάγραμμα οργανώνεται ως μια λίστα θέσεων (slots).
- Κάθε θέση αντιπροσωπεύει ενέργειες που απαιτούνται για να αναπαραχθεί ένα κομμάτι video (π.χ. ανάγνωση από τον δίσκο, μεταφορά στο δίκτυο)
- Οι ενέργειες αυτές συμβαίνουν σε χρόνο το πολύ t (block service time - **χρόνος εξυπηρέτησης κομματιού video**)
- Κάθε κατειλημμένη θέση (slot) αντιπροσωπεύει έναν πελάτη που λαμβάνει ροή δεδομένων video.
- Ο **χρόνος αναπαραγωγής κομματιού video (block play time)** T είναι ο χρόνος αναπαραγωγής του κομματιού στον υπολογιστή του πελάτη.
- Το Tiger διατηρεί διάστημα T μεταξύ των χρόνων παράδοσης των κομματιών κάθε ροής (με μικρή διακύμανση ανάλογη της δυνατότητας αποθήκευσης του πελάτη)
- Κάθε νεογνός διατηρεί ένα δείκτη στο χρονοδιάγραμμα για κάθε δίσκο του



Tiger schedule

Figure 15.11



Stream capacity of a disk = T/t (typically ~ 5)

Stream capacity of a cub with n disks = $n \times T/t$

Viewer state:

Network address of client
FileID for current movie
Number of next block
Viewer's next play slot

Cub algorithm:

- in time t {
1. Read the next block into buffer storage at the Cub.
 2. Packetize the block and deliver it to the Cub's ATM network controller with the address of the client computer.
 3. Update viewer state in the schedule to show the new next block and play sequence number and pass the updated slot to the next Cub.
 4. Clients buffer blocks and schedule their display on screen.

Χρονοπρογραμματισμός στο Tiger



- Ένας δίσκος μπορεί να διαχειρισθεί την εργασία που απαιτείται για την ταυτόχρονη εξυπηρέτηση T/t ροών.
- Για τις συνηθισμένες τιμές των T και t , $T/t > 4$
- Επομένως, ένα σύστημα Tiger με 5 νεογνούς και 3 συνδεδεμένους δίσκους σε κάθε νεογνό, μπορεί να διανείμει ταυτόχρονα > 60 ροές video

Tiger performance and scalability



1994 measurements:

- 5 x cubs: 133 MHz Pentium Win NT, 3 x 2Gb disks each, ATM network.
- supported streaming movies to 68 clients simultaneously without lost frames.
- with one cub down, frame loss rate 0.02%

1997 measurements:

- 14 x cubs: 4 disks each, ATM network
- supported streaming 2 Mbps movies to 602 clients simultaneously with loss rate of $< .01\%$
- with one cub failed, loss rate $< .04\%$

The designers suggested that Tiger could be scaled to 1000 cubs supporting 30,000 clients.