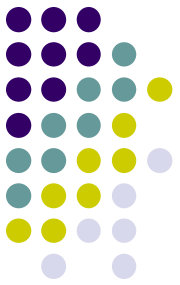


Κατανεμημένα Συστήματα

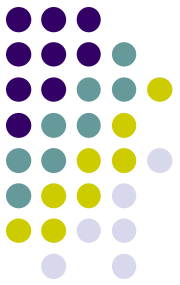
Μάθημα #3



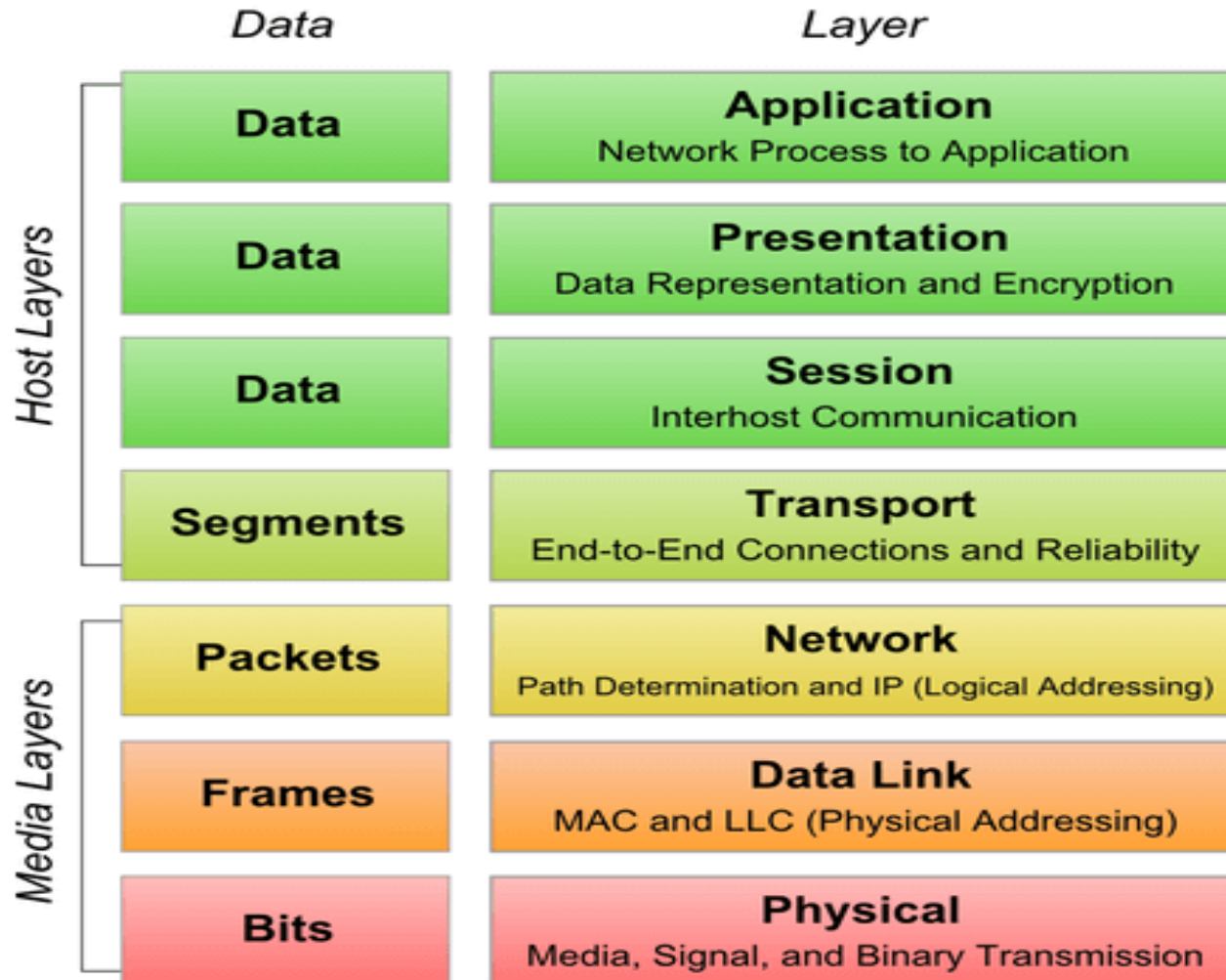
Επικοινωνία



- Στα ΚΣ όλη η επικοινωνία βασίζεται στην ανταλλαγή μηνυμάτων χαμηλού επιπέδου
- Απαιτείται συμφωνία φόρμας μηνυμάτων
- Πρότυπο
 - Στόχος τυποποίησης = Συνεργασία μεταξύ κόμβων από διαφορετικούς κατασκευαστές
 - Πρότυπο Αναφοράς για τη Διασύνδεση Ανοικτών Συστημάτων (International Standards Organization Open System Interconnection Reference Model - ISO OSI) ή Πρότυπο OSI
 - 1977 International Organization for Standardization (ISO)
- Η επικοινωνία γίνεται σε διάφορα επίπεδα μέσω πρωτοκόλλων
- Το πρωτόκολλο κάθε επιπέδου μπορεί να αλλάζει ανεξάρτητα από τα άλλα επίπεδα



OSI Model

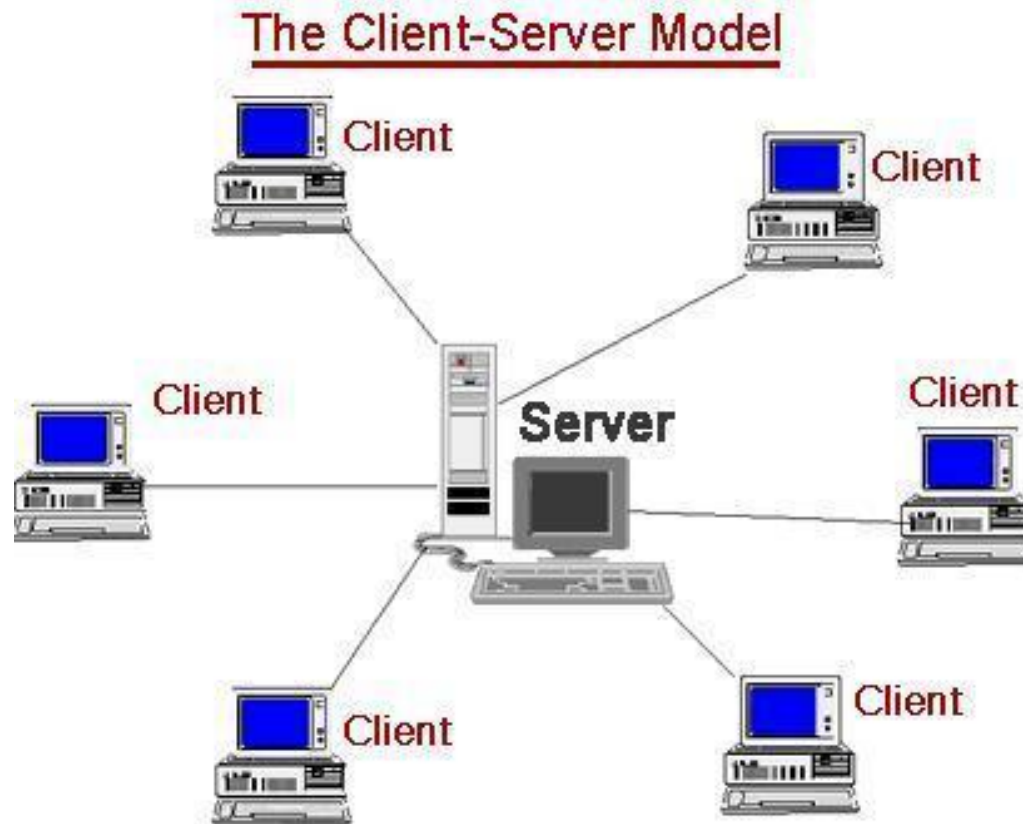
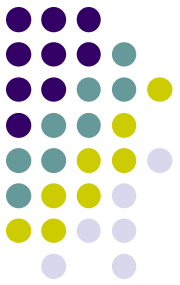




Το Μοντέλο Πελάτη-Διακομιστή

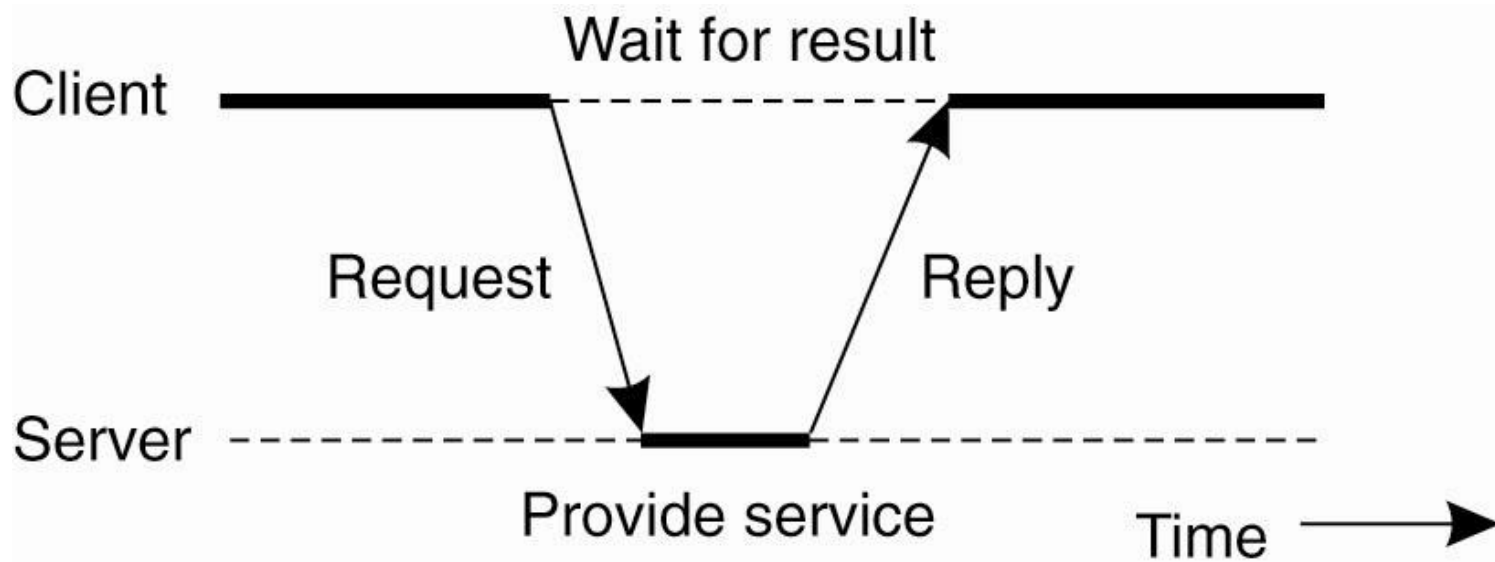
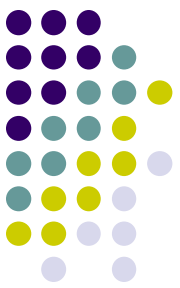
- Ένα σύστημα client-server είναι ένα σύστημα στο οποίο το δίκτυο ενώνει διάφορους υπολογιστικούς πόρους, ώστε οι clients (front end) να μπορούν να ζητούν υπηρεσίες από έναν server (back end) που προσφέρει πληροφορίες ή πρόσθετη υπολογιστική ισχύ
- Το client-server μοντέλο βασίζεται, συνήθως, σε ένα απλό πρωτόκολλο αίτησης/απάντησης (request/reply).
 - Ο πελάτης στέλνει ένα μήνυμα (αίτηση) ζητώντας από τον εξυπηρετητή κάποια υπηρεσία.
 - Ο εξυπηρετητής ενεργοποιείται άμεσα ή προσθέτει την αίτηση σε μια ουρά. Τελικά αποκρίνεται εκτελώντας μια σειρά από ενέργειες και επιστρέφει μια απάντηση, συνήθως με τα δεδομένα που ζητήθηκαν ή ένα μήνυμα λάθους.

Το Μοντέλο Πελάτη-Διακομιστή



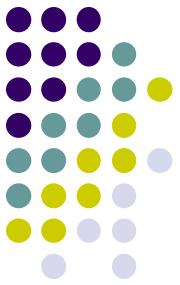
Οργάνωση Κατανεμημένων Συστημάτων

– Το Μοντέλο Πελάτη-Διακομιστή



Γενική αλληλεπίδραση μεταξύ διεργασίας-πελάτη και διεργασίας διακομιστή

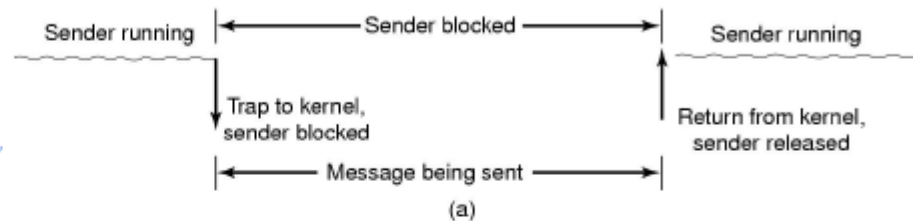
Βασική επικοινωνία στο μοντέλο πελάτη-διακομιστή (π.χ. μέσω sockets)



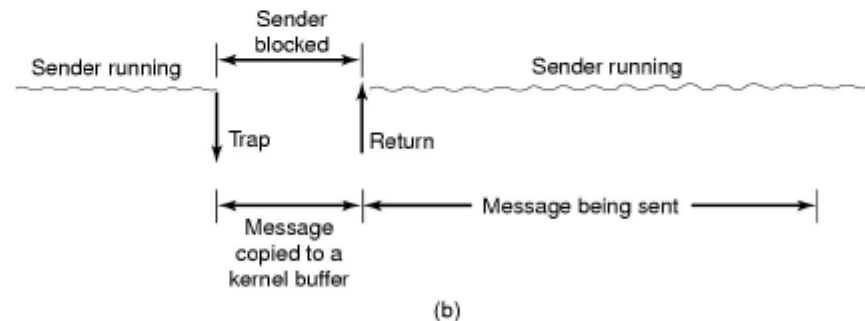
- Βασικών απαιτήσεων calls για αποστολή και παραλαβή δεδομένων (send/receive)

- Blocking και Non-blocking calls

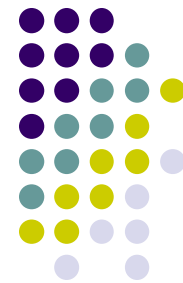
(a) Blocking send call



(b) Nonblocking send call

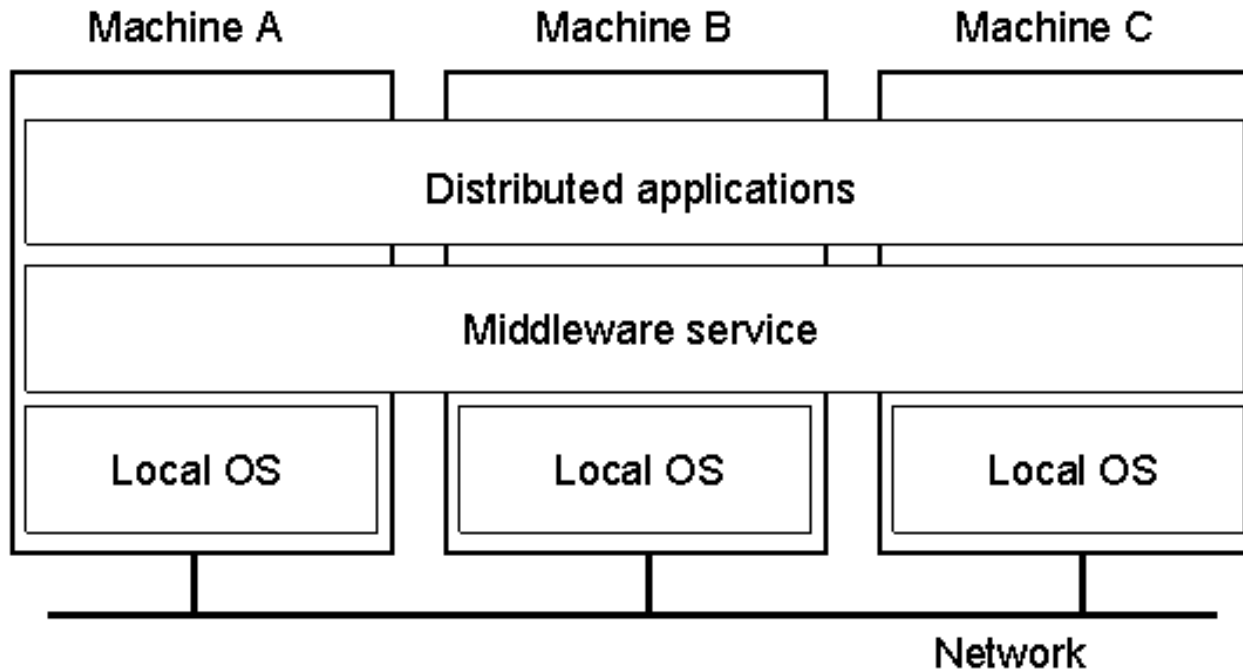
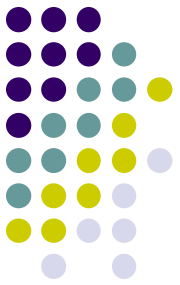


Middleware – Ενδιάμεσο Λογισμικό



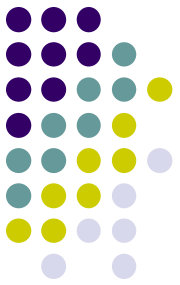
- Προκειμένου να υποστηριχτούν διαφορετικά είδη υπολογιστών και δικτύων και παράλληλα να προσφέρουν μια ομοιόμορφη εικόνα στους τελικούς χρήστες, τα ΚΣ συνήθως αποτελούνται από ένα επίπεδο λογισμικού, το **middleware**
- Το λογισμικό αυτό τοποθετείται λογικά μεταξύ ενός υψηλότερου επιπέδου που αποτελείται από τους χρήστες και τις εφαρμογές και ενός χαμηλότερου επιπέδου που αποτελείται από λειτουργικά συστήματα

Middleware



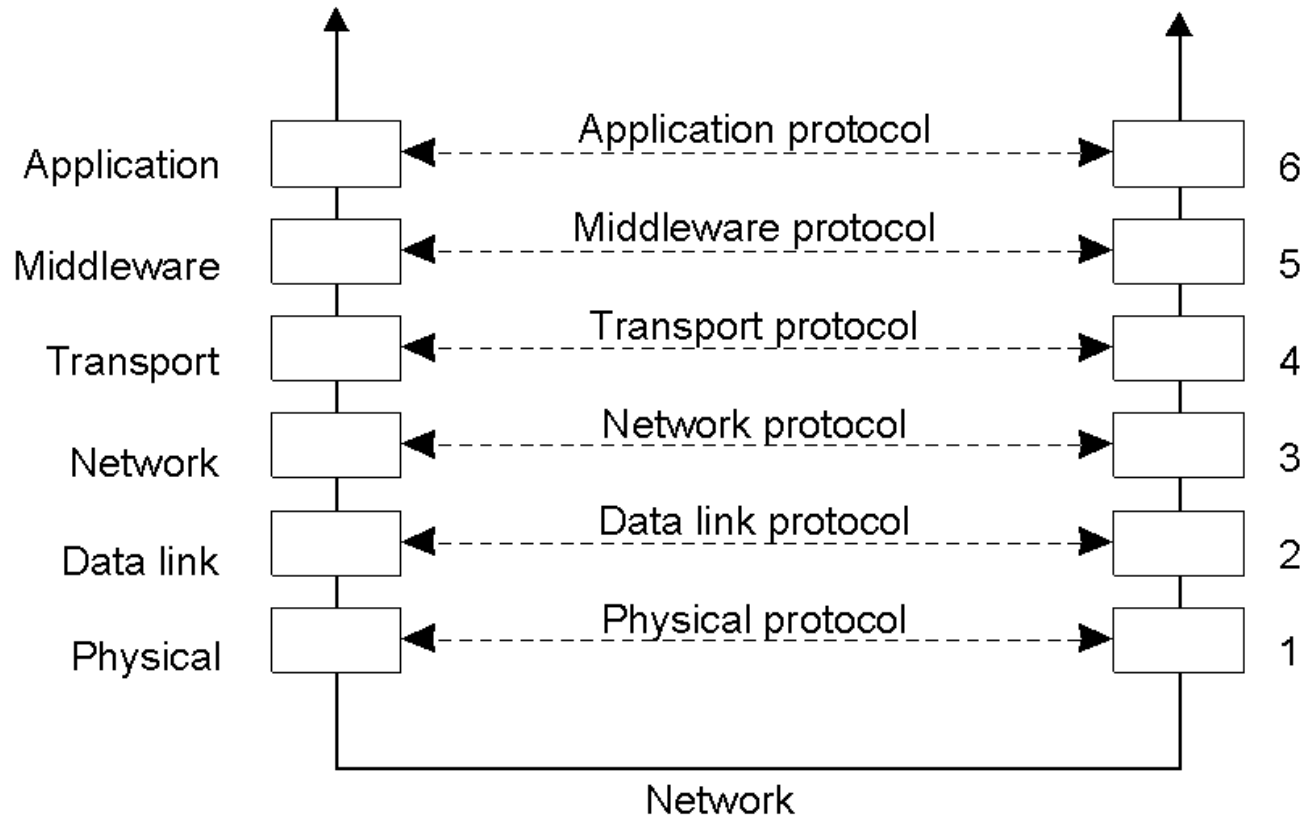
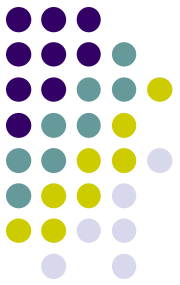
Παράδειγμα ΚΣ οργανωμένου ως middleware, το επίπεδο του οποίου καλύπτει πολλά διαφορετικά υπολογιστικά συστήματα

Μοντέλα Ενδιάμεσου Λογισμικού



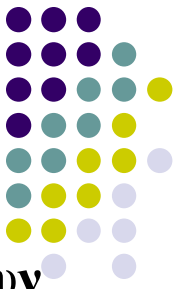
- Ενδιάμεσο λογισμικό που βασίζεται στα **ΚΣ αρχείων (NFS, HDFS κλπ)**: Η επικοινωνία ανάγεται στην προσπέλαση των διαμοιραζόμενων αρχείων.
- Ενδιάμεσο λογισμικό που βασίζεται στις **Κλήσεις Απομακρυσμένων Διαδικασιών (RPC)**: Μια διεργασία μπορεί να καλεί διαδικασία που υλοποιείται σε απομακρυσμένη μηχανή. Η καλούσα διεργασία δεν αντιλαμβάνεται ότι πραγματοποιήθηκε επικοινωνία μέσω δικτύου.
- Ενδιάμεσο λογισμικό που βασίζεται στα **Κατανεμημένα Αντικείμενα (RMI, CORBA κ.α.)**: Κάθε αντικείμενο τοποθετείται σε μία μηχανή και οι μέθοδοι που υλοποιεί γίνονται διαθέσιμοι σε άλλες μηχανές. Όταν μια διεργασία καλεί μια μέθοδο, στέλνεται μήνυμα στο αντικείμενο το οποίο εκτελεί τη μέθοδο και στέλνει πίσω το αποτέλεσμα.

Πρωτόκολλα ενδιάμεσου λογισμικού



An adapted reference model for networked communication.

Πρωτόκολλα ενδιάμεσου λογισμικού



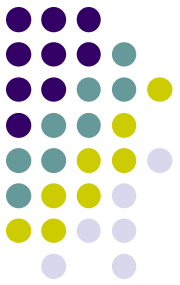
Τα πρωτόκολλα επικοινωνίας του ενδιάμεσου λογισμικού υποστηρίζουν ποικιλία υπηρεσιών ενδιάμεσου λογισμικού:

- Πρωτόκολλα πιστοποίησης ταυτότητας (Authentication)
- Πρωτόκολλα εξουσιοδότησης πρόσβασης (Authorization) σε πιστοποιημένους χρήστες
- Πρωτόκολλα καταναεμημένης δέσμευσης (Distributed commit) ή ατομικότητας
- Πρωτόκολλα καταναεμημένου κλειδώματος (προστασία πόρων από ταυτόχρονη πρόσβαση)

Τα πρωτόκολλα επικοινωνίας του ενδιάμεσου λογισμικού υποστηρίζουν υπηρεσίες επικοινωνίας υψηλού επιπέδου:

- κλήση απομακρυσμένων διαδικασιών
- κλήση απομακρυσμένων αντικειμένων

Remote Procedure Call (RPC)



- Κλήση Απομακρυσμένων Διαδικασιών

Όταν η επικοινωνία γίνεται με ανταλλαγή μηνυμάτων π.χ. κλήσεις send/receive χάνεται η διαφάνεια του συστήματος

Birrell and Nelson (1984)

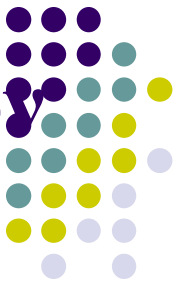
“Να επιτρέπεται στα προγράμματα να καλούν διαδικασίες που βρίσκονται σε άλλες μηχανές.”

Η μέθοδος αυτή λέγεται **Remote Procedure Call - (RPC)**

Προβλήματα

- Two machine architectures may not (or need not) be identical.
- Each machine has a different address space.
- How are parameters (of different, possibly very complex, types) passed to/from a remote procedure?

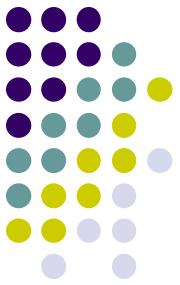
Συμβατική (Τοπική) Κλήση Διαδικασιών



Μεταβίβαση παραμέτρων

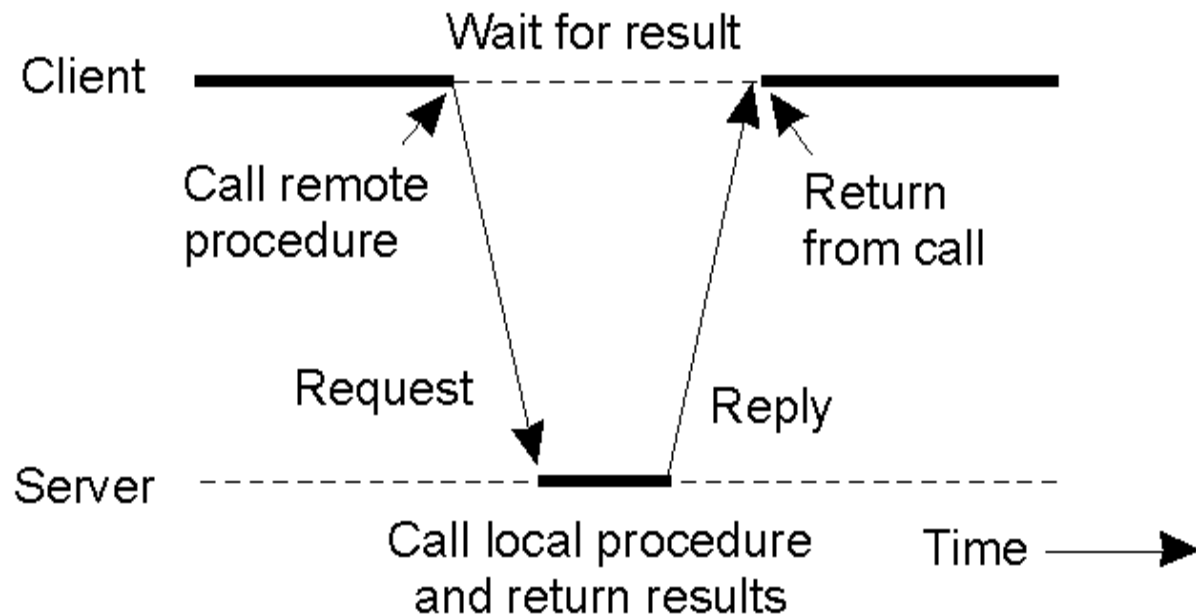
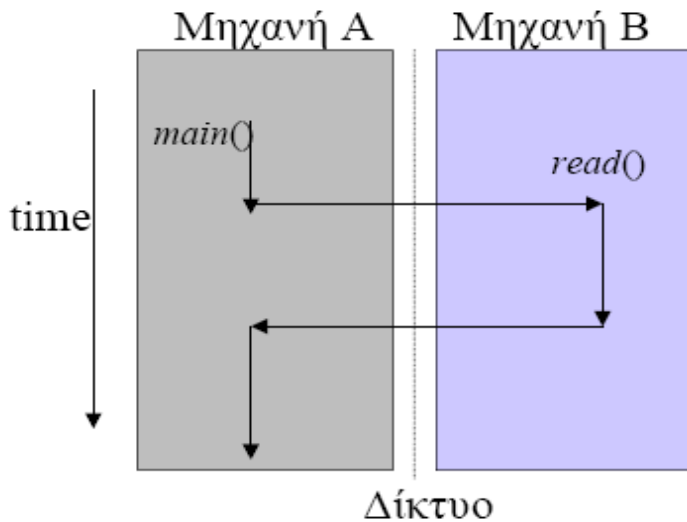
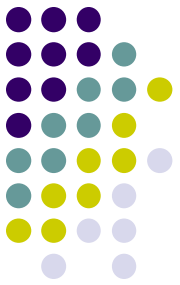
- **Μεταβίβαση με τιμή**
 - Αντιγραφή στη στοίβα της πραγματικής παραμέτρου
 - Καλούμενη διαδικασία: η παράμετρος είναι τοπική μεταβλητή
 - οποιεσδήποτε αλλαγές δεν επηρεάζουν την αρχική τιμή της παραμέτρου
- **Μεταβίβαση με αναφορά**
 - Δείκτης στην πραγματική παράμετρο
 - Σύνδεση με μεταβλητή του καλούντα

Επικοινωνία με χρήση απομακρυσμένων διαδικασιών



- Τα προγράμματα μπορούν να καλούν διαδικασίες που εκτελούνται σε διαφορετικές μηχανές (Birrell & Nelson 1984)
- Λογική αφαίρεση υψηλότερου επιπέδου
- Παρόμοια με κλήσεις τοπικών διαδικασιών
- Όταν μία διαδικασία στη μηχανή A καλέσει μία διαδικασία στη μηχανή B, η εκτέλεση της διεργασίας στην A αναστέλλεται και η εκτέλεση της καλούμενης διαδικασίας συνεχίζεται στη μηχανή B
- Πληροφορίες δίνονται από την καλούσα και την καλούμενη με τη μορφή παραμέτρων
- Απόκρυψη της πραγματικής επικοινωνίας - Δεν λαμβάνεται γνώση της μεταβίβασης των μηνυμάτων και της λειτουργίας εισόδου/εξόδου

Κλήση απομακρυσμένων διαδικασιών



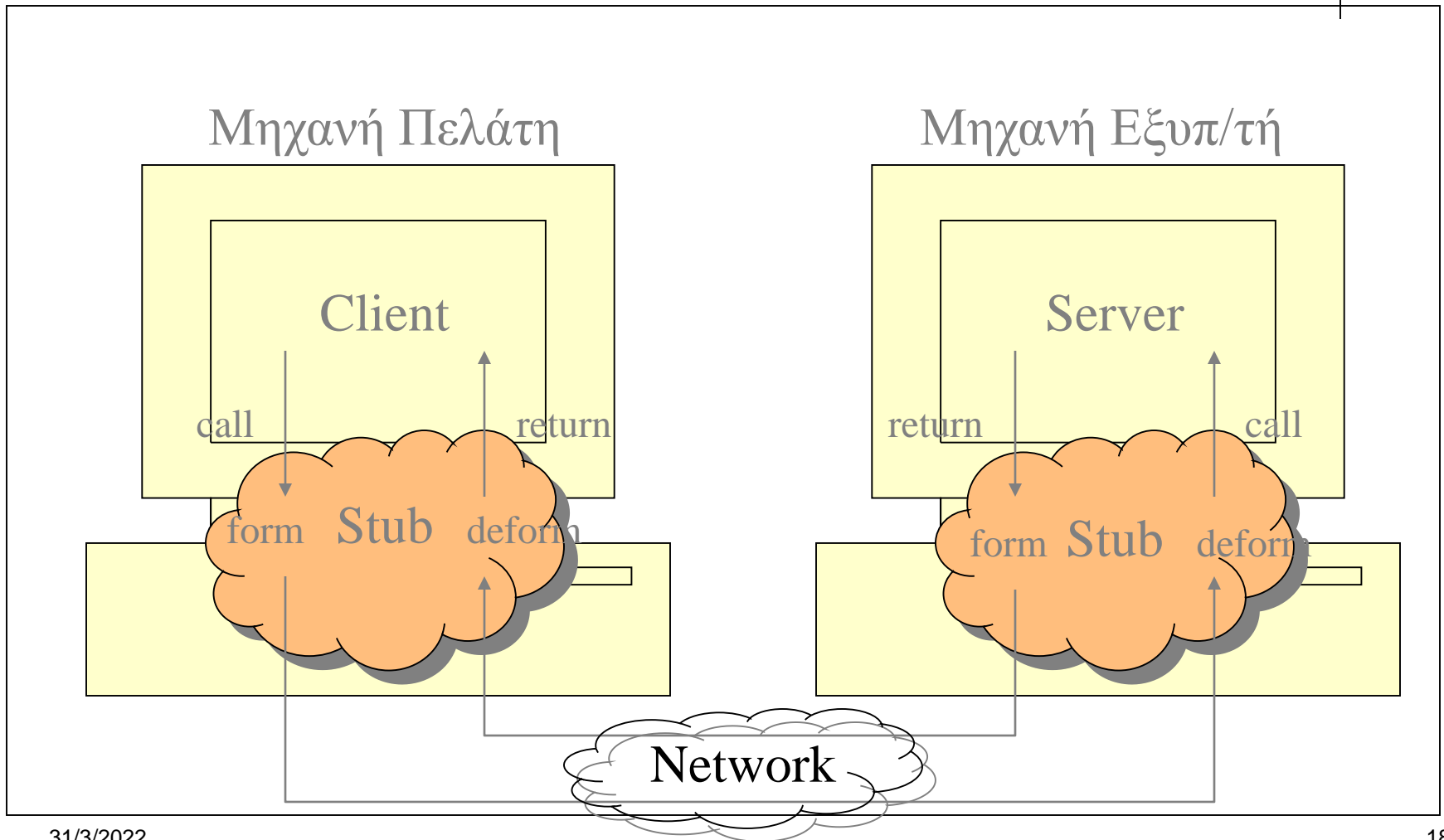
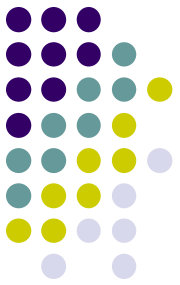


Λειτουργία μηχανισμού RPC

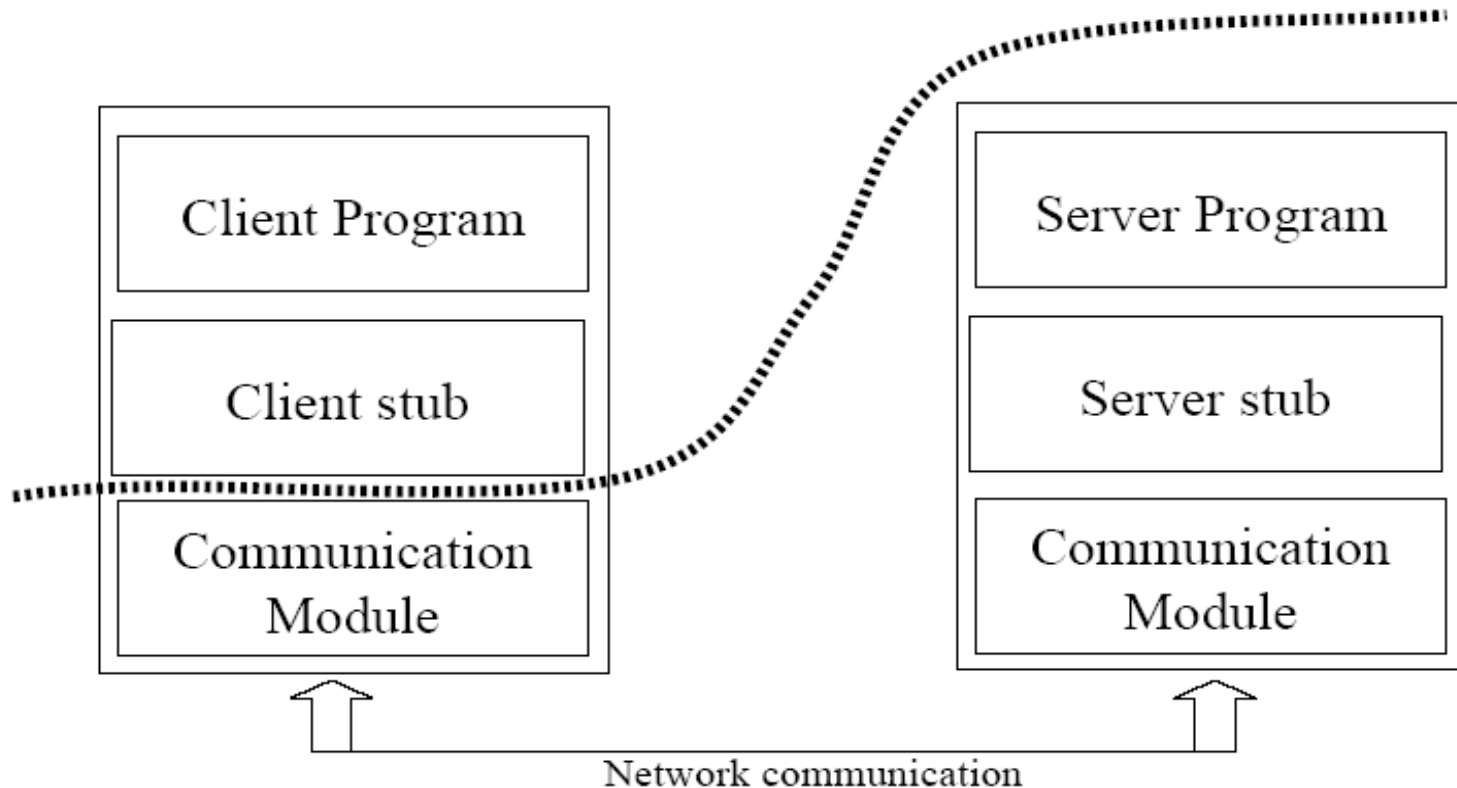
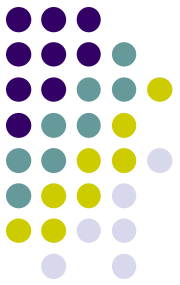
Η διαφάνεια της λειτουργίας του μηχανισμού RPC επιτυγχάνεται με την χρήση στελεχών-υποκατάστατων (stubs) client & server της διαδικασίας που καλείται

- Υλοποίηση της απομακρυσμένης κλήσης με 2 τοπικές κλήσεις
 - Κλήση στελέχους πελάτη τοπικά από πελάτη
 - Κλήση διαδικασίας τοπικά από το στέλεχος διακομιστή
- Στέλεχος πελάτη (client stub)
 - Ίδια κλητική ακολουθία με διαδικασία
 - Εκδοχή της απομακρυσμένης διαδικασίας στον πελάτη
- Στέλεχος διακομιστή (server stub)
 - Εκτελεί την κλήση της διαδικασίας
 - Μετατρέπει τις εισερχόμενες κλήσεις σε τοπικές κλήσεις

Κλήση Απομακρυσμένων Διαδικασιών



Client και Server

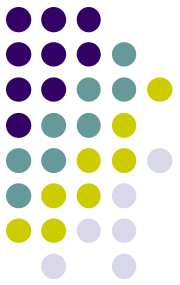


Client program only communicates with the client stub



Βήματα για απομακρυσμένη κλήση

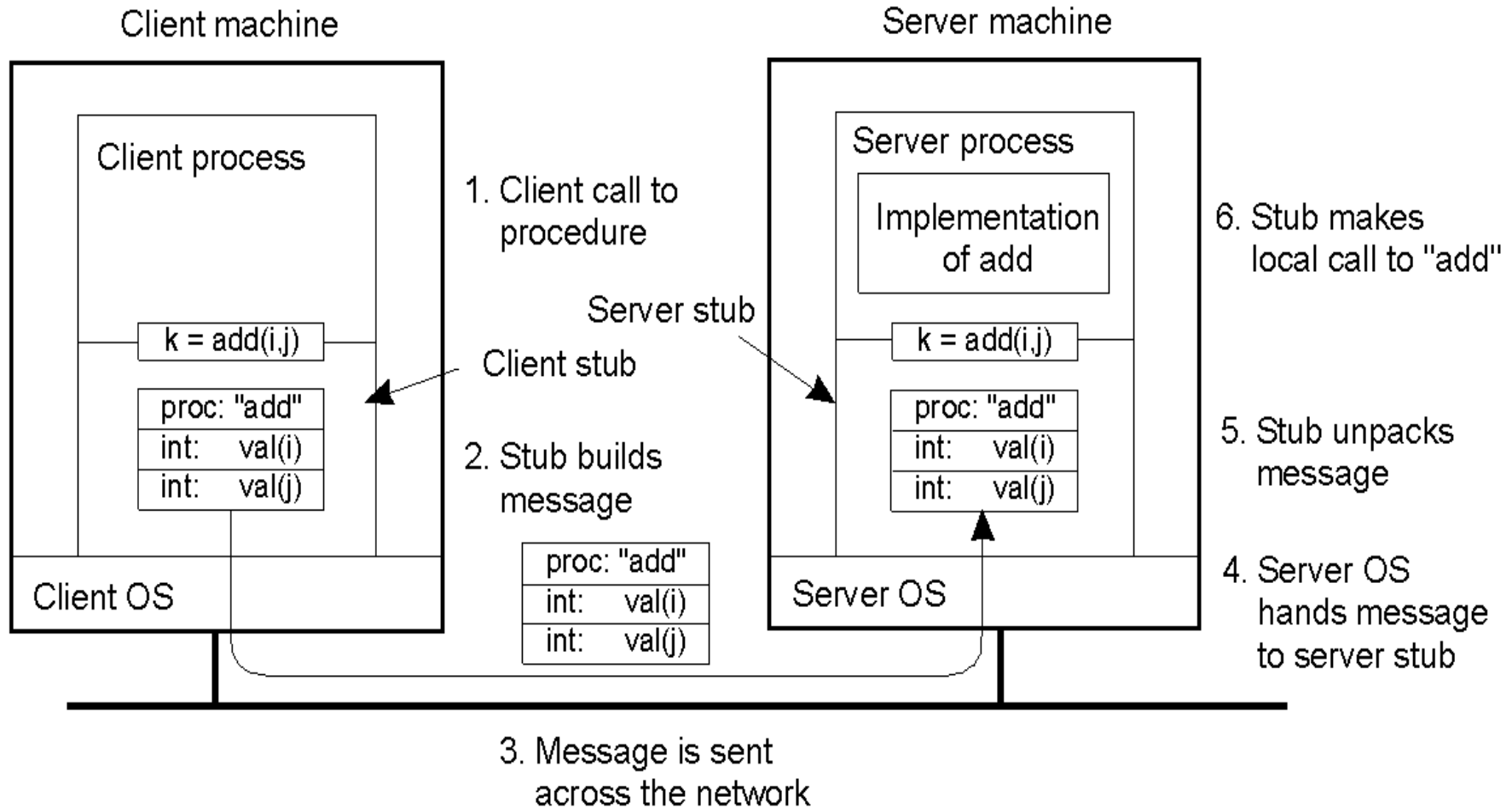
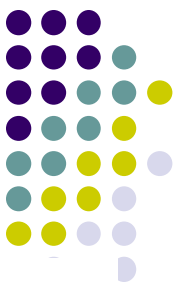
- Η διαδικασία του πελάτη καλεί το στέλεχος πελάτη
- Το στέλεχος πελάτη σχηματίζει ένα μήνυμα και καλεί το τοπικό ΛΣ που στέλνει το μήνυμα στο απομακρυσμένο ΛΣ
- Το απομακρυσμένο ΛΣ παραδίδει το μήνυμα στο στέλεχος διακομιστή
- Το στέλεχος διακομιστή ξεπακετάρει τις παραμέτρους από το μήνυμα και καλεί τον διακομιστή για να το εκτελέσει
- Ο διακομιστής επιστρέφει τα αποτελέσματα στο στέλεχος του διακομιστή, το οποίο σχηματίζει μήνυμα με τα αποτελέσματα και καλεί το τοπικό ΛΣ
- Το ΛΣ του διακομιστή στέλνει το μήνυμα στο ΛΣ του πελάτη
- Το ΛΣ του πελάτη παραδίδει το μήνυμα στο στέλεχος πελάτη που ξεπακετάρει και επιστρέφει τα αποτελέσματα στον πελάτη



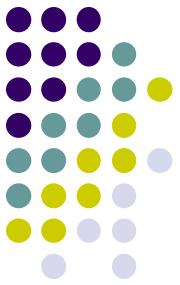
Μεταβίβαση Παραμέτρων

Πληροφορίες δίνονται από την καλούσα και την καλούμενη διεργασία με τη μορφή παραμέτρων

Μεταβίβαση παραμέτρων κατ' αξία



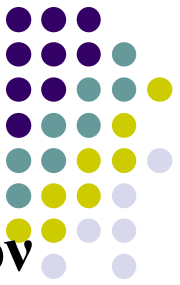
Προβλήματα



Μεταβίβαση παραμέτρων

- Αν ο πελάτης και ο εξυπηρετητής έχουν διαφορετικές αρχιτεκτονικές, χρειάζεται συμφωνία στο format μεταξύ πελάτη και εξυπηρετητή

Μεταβίβαση παραμέτρων κατ' αναφορά

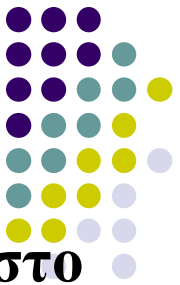


- Ένας δείκτης έχει νόημα μόνο μέσα στο χώρο διευθύνσεων της διεργασίας όπου χρησιμοποιείται.
- Επομένως δεν μπορούν να μεταβιβαστούν δείκτες.

Τι μπορεί να γίνει?

Ας υποθέσουμε ότι έχουμε ως παράμετρο της διαδικασίας, έναν δείκτη σε έναν πίνακα χαρακτήρων (buf) και ότι γνωρίζουμε το μέγεθος του πίνακα.

Τότε μπορεί να λάβει χώρα η ακόλουθη διαδικασία η οποία αντικαθιστά την μεταβίβαση παραμέτρων κατ' αναφορά με αντιγραφή-επαναφορά.

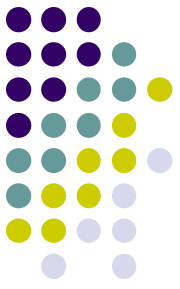


Μεταβίβαση παραμέτρων με αντιγραφή-επαναφορά.

1. Ο πίνακας αντιγράφεται μέσα στο μήνυμα και στέλνεται στο διακομιστή.
2. Το στέλεχος διακομιστή καλεί το διακομιστή με παράμετρο ένα δείκτη σε αυτόν τον πίνακα (ο δείκτης θα έχει διαφορετική τιμή από τον αρχικό). Οι αλλαγές που κάνει ο διακομιστής, επηρεάζουν τον χώρο προσωρινής αποθήκευσης του μηνύματος στο στέλεχος του διακομιστή.
3. Όταν ο διακομιστής τελειώσει, το αρχικό μήνυμα στέλνεται στο στέλεχος πελάτη.
4. Το στέλεχος πελάτη αντιγράφει τον πίνακα πίσω στον πελάτη.

Βελτιστοποίηση: η μία αντιγραφή μπορεί να παραλειφθεί.

- Αν ο πίνακας είναι είσοδος για το διακομιστή δε χρειάζεται να σταλεί στο στέλεχος πελάτη.
- Αν είναι έξοδος δε χρειάζεται να σταλεί αρχικά στο διακομιστή.



Μεταβίβαση παραμέτρων με αντιγραφή-επαναφορά

- Είναι δυνατός ο χειρισμός δεικτών σε πίνακες με γνωστό μέγεθος.
- Δεν είναι δυνατός ο χειρισμός δεικτών σε μια οποιαδήποτε δομή δεδομένων. **Για το χειρισμό τέτοιων δεικτών επιχειρούνται λύσεις όπως:** μεταβιβάζεται ο δείκτης στο στέλεχος διακομιστή και στέλνεται πίσω στον πελάτη αίτηση για την αποστολή των αναφερόμενων δεδομένων.

Προσδιορισμός παραμέτρων



```
foobar( char x; float y; int z[5] )  
{  
    ....  
}
```

(a)

foobar's local variables	
	x
y	
5	
z[0]	
z[1]	
z[2]	
z[3]	
z[4]	

(b)

(α) μια διαδικασία (β) το αντίστοιχο μήνυμα

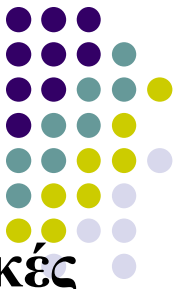
Ορισμός πρωτοκόλλου RPC



Και οι δύο πλευρές μιας απομακρυσμένης κλήσης θα πρέπει να ακολουθούν το ίδιο πρωτόκολλο.

- **Ορισμός μορφής μηνυμάτων – προσδιορισμός παραμέτρων**
- **Ο πελάτης και ο διακομιστής θα πρέπει να συμφωνούν στην αναπαράσταση των απλών δομών δεδομένων (int, float, char) ώστε τα μηνύματα να ερμηνεύονται χωρίς ασάφειες.**
- **Ο πελάτης και ο διακομιστής θα πρέπει να συμφωνούν για το πώς θα γίνει η ίδια η ανταλλαγή μηνυμάτων (π.χ. χρήση TCP/IP ή UDP σε συνδυασμό με ένα μηχανισμό ελέγχου ασφαλείας που υλοποιούν οι ίδιοι στο πλαίσιο του πρωτοκόλλου RPC).**

Υλοποίηση στελεχών



- Τα στελέχη για το ίδιο πρωτόκολλο αλλά για διαφορετικές διαδικασίες διαφέρουν μόνο στη διασύνδεσή τους με τις εφαρμογές.
- Μία διασύνδεση αποτελείται από μια συλλογή διαδικασιών που μπορούν να καλούνται από έναν πελάτη και υλοποιούνται από έναν διακομιστή.
- Οι διασυνδέσεις ορίζονται σε μια γλώσσα ορισμού διασυνδέσεων (Interface Definition Language – IDL)
- Μία διασύνδεση μεταγλωττίζεται σε ένα στέλεχος πελάτη και ένα στέλεχος διακομιστή.

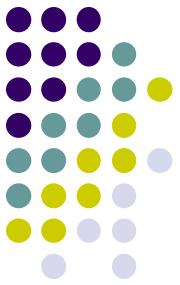
Όλα τα συστήματα ενδιάμεσου λογισμικού που βασίζονται σε **RPC** παρέχουν μια γλώσσα **IDL** για την υποστήριξη της ανάπτυξης εφαρμογών.

Επεκτάσεις του αρχικού μοντέλου RPC



- What if doing RPC to a process on the same machine?
 - Can use shared memory to optimize?
- Does RPC give you concurrency? That is, can you have the caller and callee executing at the same time?
 - Does normal procedure call provide that?
- Does RPC allow you to overlap communication with computation?
 - Suppose you are sending an array of 1 GB.

Ασύγχρονες RPCs

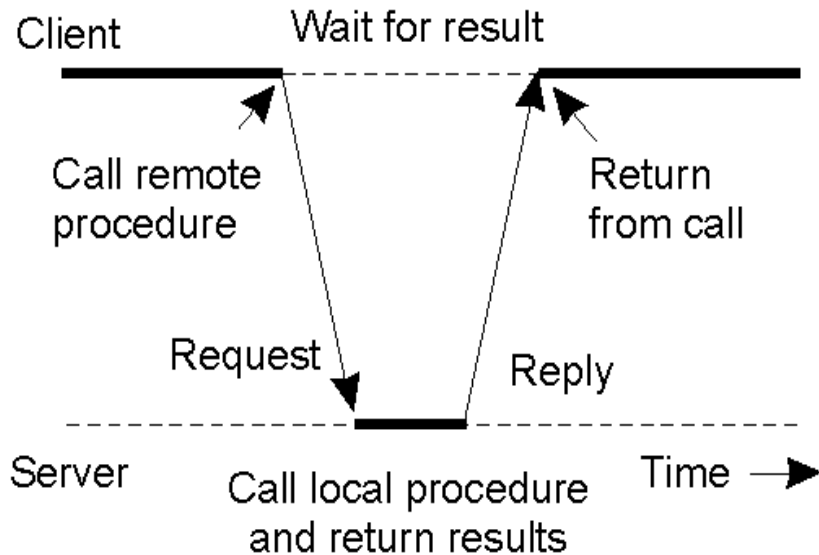


- **Συμβατικές RPCs:** ο πελάτης μπλοκάρεται μέχρι να επιστραφεί απάντηση

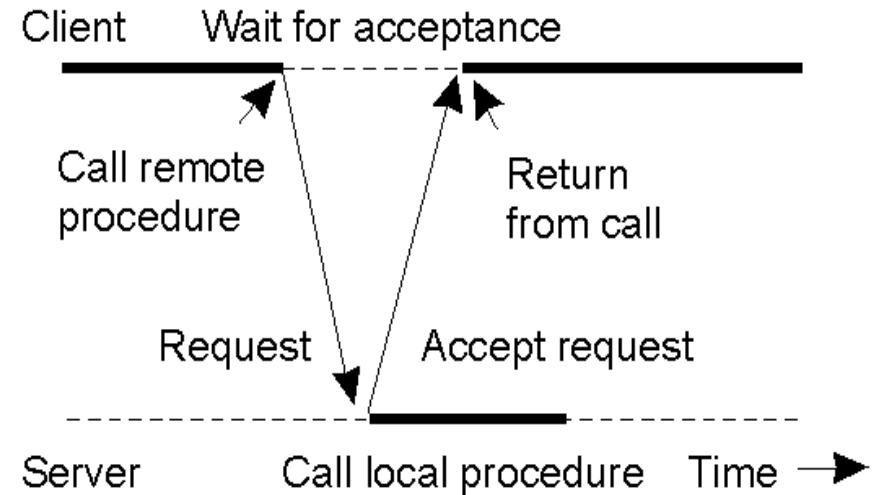
Αν δεν χρειάζεται απάντηση (π.χ., ενημέρωση στοιχείων, εκκίνηση απομακρυσμένων διαδικασιών, μεταφορά χρημάτων)?

- **Ασύγχρονες RPCs:** ο πελάτης συνεχίζει αμέσως μετά την παραλαβή επιβεβαίωσης αποδοχής της αίτησης από τον διακομιστή

Ασύγχρονες RPCs



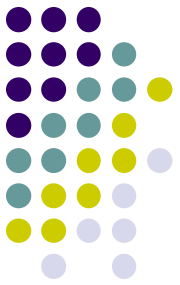
(a)



(b)

- a) The interconnection between client and server in a traditional RPC
- b) The interaction using asynchronous RPC

Ασύγχρονες RPCs



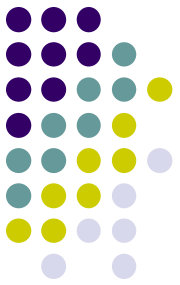
Αν πρόκειται να επιστραφεί απάντηση αλλά ο πελάτης δεν είναι διατεθειμένος να περιμένει:

Καθυστερημένη σύγχρονη RPC: η επικοινωνία επιτυγχάνεται με δύο ασύγχρονες RPCs.

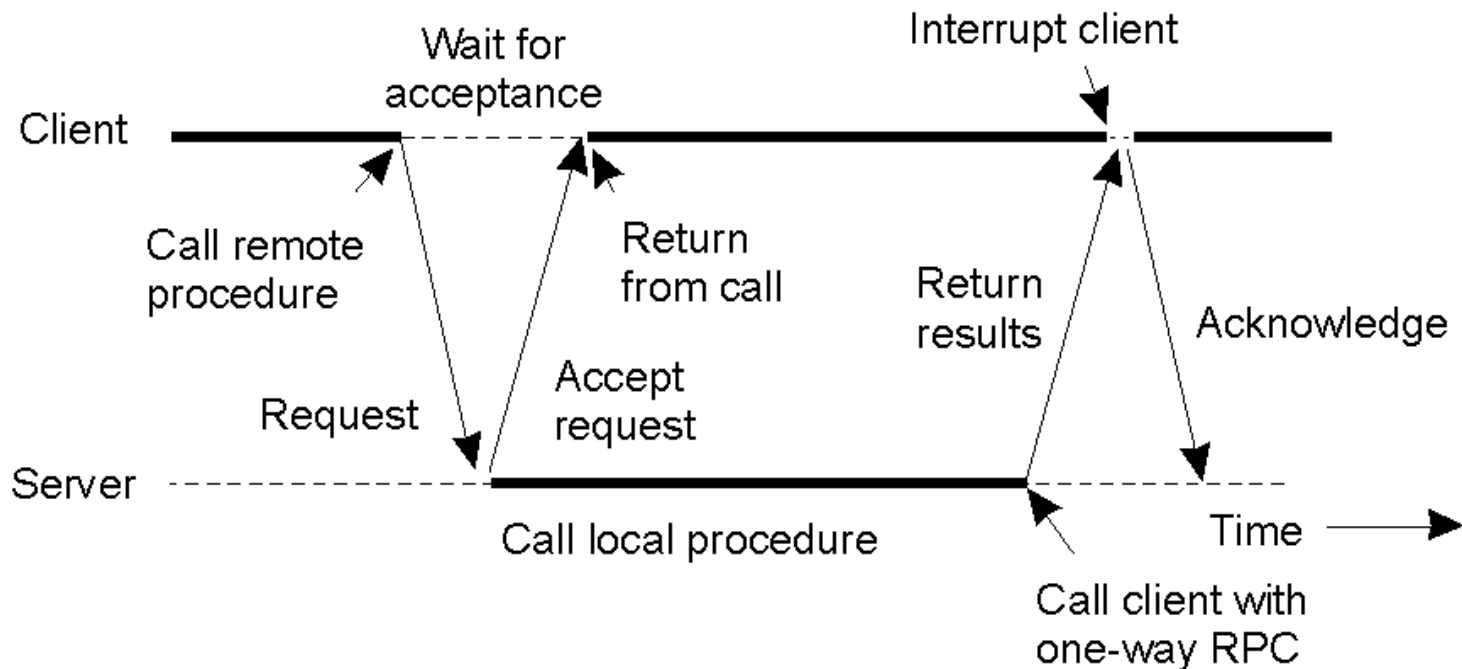
Μονόδρομη RPC: ο πελάτης δεν περιμένει για επιβεβαίωση της αποδοχής της αίτησης από τον διακομιστή

Πρόβλημα αξιοπιστίας: ο πελάτης δε γνωρίζει με βεβαιότητα ότι το μήνυμά του έχει παραληφθεί και η αίτηση θα διεκπεραιωθεί

Ασύγχρονες RPCs



Καθυστερημένη σύγχρονη RPC: η επικοινωνία επιτυγχάνεται με δύο ασύγχρονες RPCs.

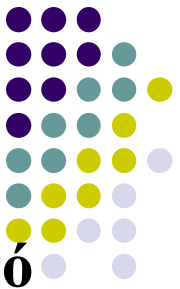


Διαφάνεια RPC



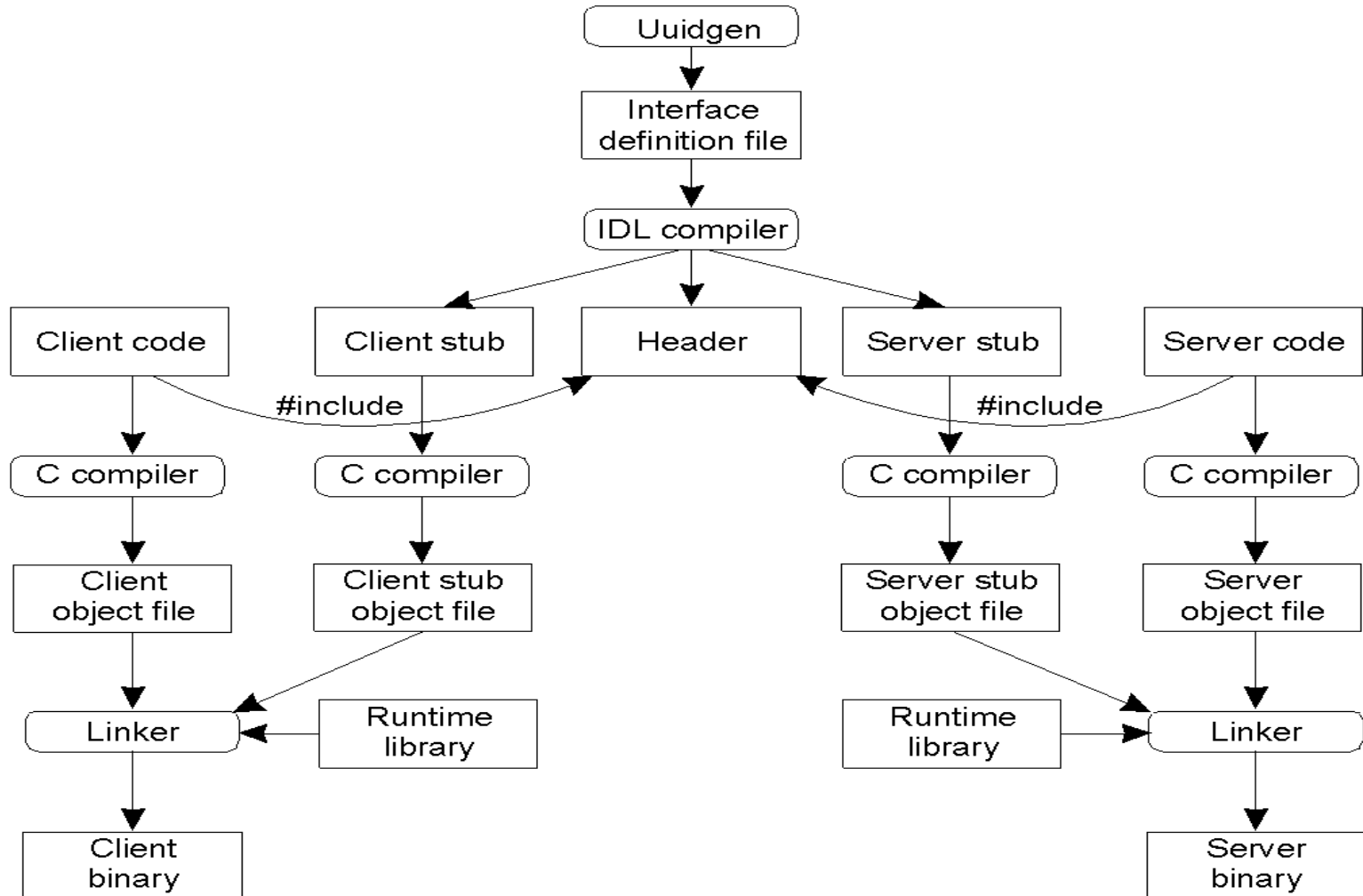
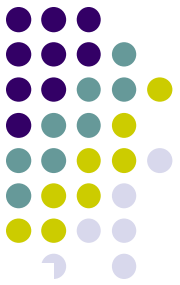
- **Συντακτική (syntactic) και σημασιολογική (semantic) διαφάνεια**
- **Η συντακτική διαφάνεια δεν είναι δύσκολο να επιτευχθεί, αφού οι απομακρυσμένες κλήσεις ακολουθούν το ίδιο μοντέλο με τις τοπικές**
 - **Κλήση απομακρυσμένης διαδικασίας από μία διεργασία και αναστολή της διεργασίας μέχρι την επιστροφή της κλήσης**
 - **Πέρασμα παραμέτρων - επιστροφή αποτελεσμάτων στην κλήση**
- **Η πλήρης σημασιολογική διαφάνεια είναι δύσκολο να επιτευχθεί αφού οι απομακρυσμένες διαδικασίες**
 - **εκτελούνται σε ξεχωριστό address space και απαιτούν ειδικούς μηχανισμούς (call-by-copy/restore) αντί call by value/reference**
 - **παρουσιάζουν σφάλματα λόγω διαφορετικών μηχανών, δικτύου**
 - **χρειάζονται 10-1000 φορές περισσότερο χρόνο, γεγονός που πρέπει να ξέρει η διεργασία πελάτη**

Distributed Computing Environment (DCE RPC)

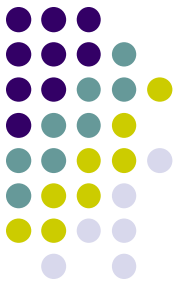


- Είναι σύστημα ενδιάμεσου λογισμικού που σχεδιάστηκε από την OSF (Open Software Foundation – Open Group).
Σχεδιάστηκε για το UNIX
- Μπορεί να εντοπίζει αυτόματα το σωστό διακομιστή και να εγκαθιστά την επικοινωνία μεταξύ πελάτη και διακομιστή (συνδυασμός – binding).
- Χειρίζεται αυτόματα μετατροπές τύπων δεδομένων, ακόμη και αν ο πελάτης και ο διακομιστής εκτελούνται σε διαφορετικές αρχιτεκτονικές
- Αποτελείται από ένα πλήθος στοιχείων (γλώσσες, βιβλιοθήκες, βοηθητικά προγράμματα) προκειμένου να είναι δυνατή η συγγραφή πελατών και διακομιστών.

Δημιουργία πελάτη και διακομιστή στο DCE RPC

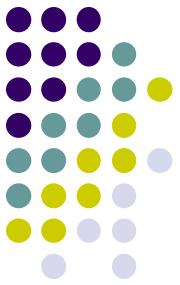


Δημιουργία πελάτη και διακομιστή στο DCE RPC



- Τα αρχεία IDL περιλαμβάνουν τις δηλώσεις των διαδικασιών, ορισμούς τύπων, δηλώσεις σταθερών και άλλες πληροφορίες για τη σωστή παράταξη των παραμέτρων.
- Κάθε αρχείο IDL περιέχει ένα καθολικά μοναδικό αναγνωριστικό για τη διασύνδεση που ορίζεται. Ο πελάτης στέλνει αυτό το αναγνωριστικό με το πρώτο μήνυμα RPC και ο διακομιστής επαληθεύει αν είναι σωστό.

Βήματα συγγραφής εφαρμογής πελάτη διακομιστή



- Κλήση προς το πρόγραμμα `uiidgen` το οποίο παράγει αρχείο IDL με περιέχει μοναδικό κωδικό
- Επεξεργασία του αρχείου IDL με τη συμπλήρωση των ονομάτων των απομακρυσμένων διαδικασιών και των παραμέτρων τους
- Η έξοδος του μεταγλωττιστή IDL αποτελείται από:
 - (α) **αρχείο κεφαλίδων** (μοναδικό αναγνωριστικό, ορισμούς τύπων, κ.λπ.)
 - (β) **στέλεχος πελάτη** (περιλαμβάνει τις διαδικασίες που είναι υπεύθυνες για τη συλλογή και συσκευασία των παραμέτρων στο εξερχόμενο μήνυμα και την κλήση του runtime system για την αποστολή του μηνύματος)
 - (γ) **στέλεχος διακομιστή** (περιλαμβάνει τις διαδικασίες που είναι υπεύθυνες για την κλήση των διαδικασιών διακομιστή που διεκπεραιώνουν την εργασία)
- Συγγραφή του κώδικα του πελάτη και του διακομιστή

Συνδυασμός (binding) πελάτη και διακομιστή



Εντοπισμός του διακομιστή (διεργασίας):

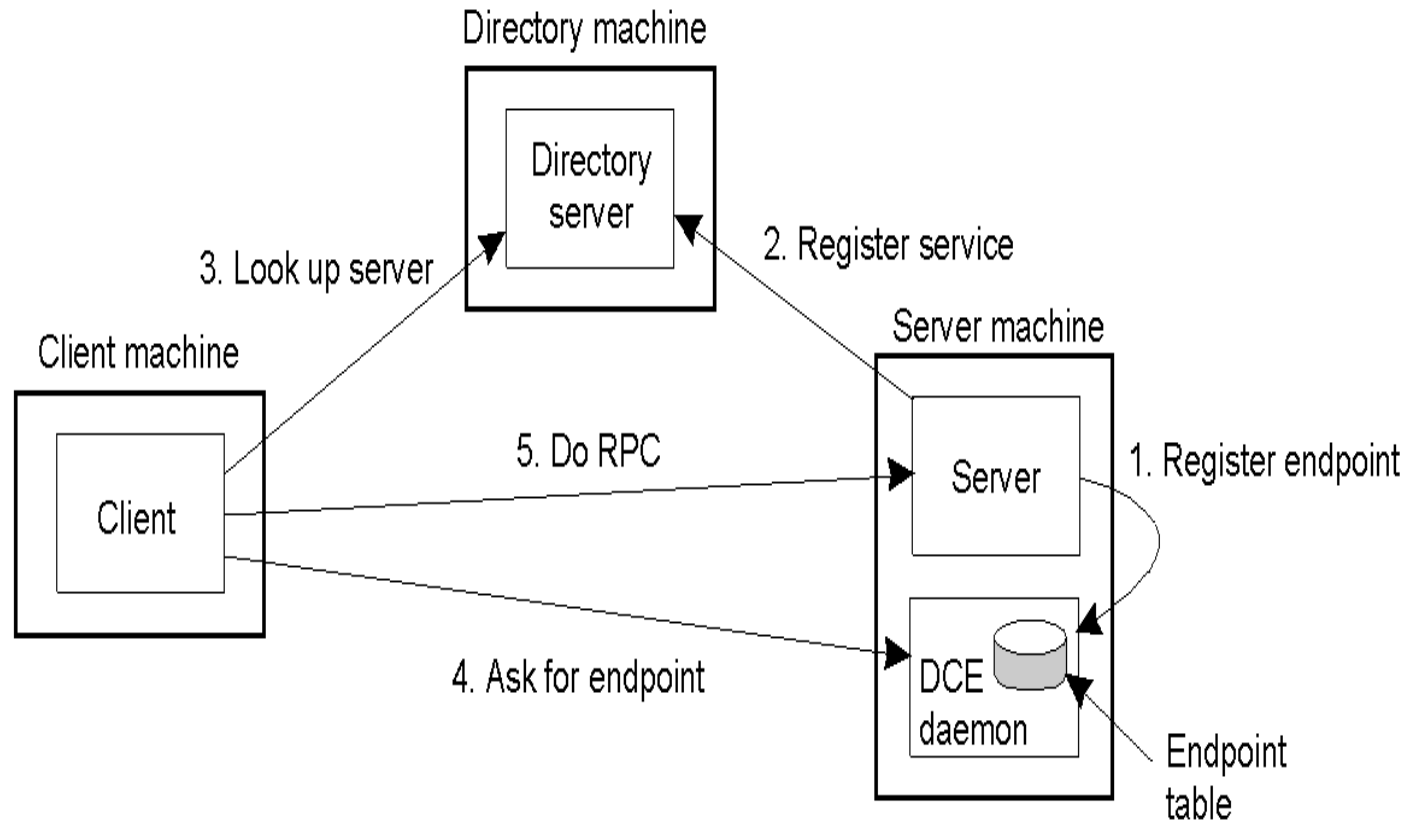
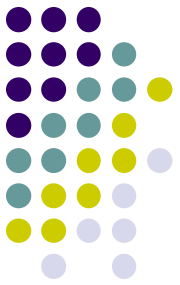
1. Εντοπισμός της μηχανής του διακομιστή μέσω της υπηρεσίας ευρετηρίου

- Η διεύθυνση της υπηρεσίας ευρετηρίου πρέπει να είναι γνωστή (π.χ. fixed ή μέσω αρχικού broadcast)
- Ο διακομιστής (διεύθυνση δικτύου και όνομα διακομιστή) καταχωρίζεται στην υπηρεσία ευρετηρίου
- Ο διακομιστής μπορεί να διαγραφεί από την υπηρεσία ευρετηρίου όταν δεν επιθυμεί να προσφέρει πλέον υπηρεσίες

2. Εντοπισμός του διακομιστή (διεργασίας) στη μηχανή διακομιστή

- Για να επικοινωνήσει ο πελάτης με έναν διακομιστή πρέπει να γνωρίζει ένα ακραίο σημείο (θύρα) στη μηχανή του διακομιστή όπου θα στείλει το μήνυμα (τα ακραία σημεία χρησιμοποιούνται από το Λ.Σ. για τη διάκριση των μηνυμάτων που απευθύνονται σε διαφορετικές διεργασίες).
- Ο Δαίμονας DCE σε κάθε διακομιστή κρατάει πίνακα ζευγών (διεργασία, ακραίο_σημείο)

Συνδυασμός (binding) πελάτη και διακομιστή στο DCE





Κλήσεις RPC στο DCE

Προεπιλογή: Καμία κλήση δεν διεκπεραιώνεται ποτέ πανω από μία φορά ακόμη και σε περιπτώσεις καταρρευσης συστήματος.

Μπορεί όμως να οριστεί στο IDL ότι αν μια κλήση αποτύχει λόγω προβλημάτων επικοινωνίας, τότε αυτή να επαναλαμβάνεται αυτόματα μέχρι να ολοκληρωθεί επιτυχώς.

Θα πρέπει η κλήση αυτή να μην τροποποιεί δεδομένα στον διακομιστή.