



Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Ηλεκτρολόγων Μηχ. και Μηχανικών Υπολογιστών
Εργαστήριο Υπολογιστικών Συστημάτων

Χρονοδρομολόγηση ΚΜΕ

(==Χρονοπρογραμματισμός)

Λειτουργικά Συστήματα Υπολογιστών
6ο Εξάμηνο, 2019-2020

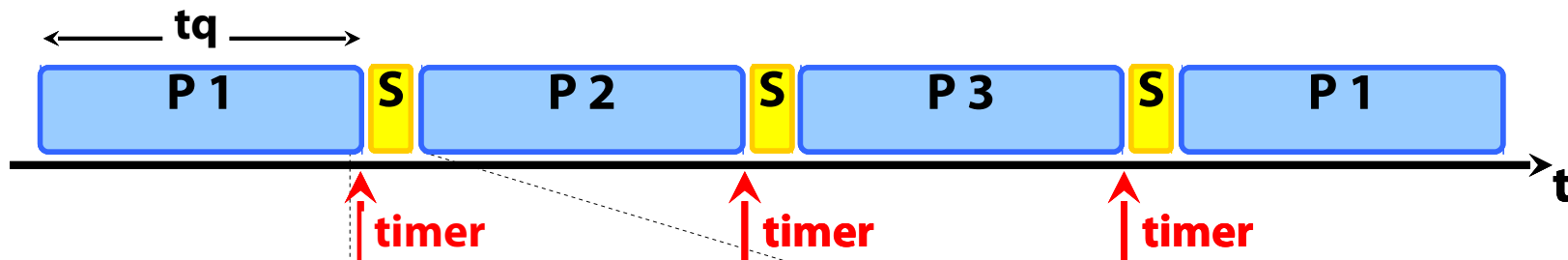
Σκελετός Παρουσίασης

- Μοντέλο Διαμοιρασμού χρόνου και Συμπεριφορά Διεργασιών
- Χρονοδρομολογητής
- Αλγόριθμοι Χρονοδρομολόγησης
- Χρονοδρομολόγηση σε Συστήματα Πολλαπλών Επεξεργαστών

Μοντέλο Διαμοιρασμού Χρόνου

- Πολλαπλές διεργασίες (P_0, P_1, P_2, \dots)
- Οι διεργασίες έχουν την (ψευδ-)αίσθηση ότι χρησιμοποιούν αποκλειστικά τον επεξεργαστή
- Ο **χρονοδρομολογητής (ΧΔ)** αναλαμβάνει:
 - Την επιλογή της διεργασίας που θα χρησιμοποιήσει τον επεξεργαστή
 - Την αλλαγή της διεργασίας που εκτελείται στον επεξεργαστή (*context switch*)

Διαμοιρασμός χρόνου με διακοπές χρονιστή



- Επιλογή επόμενης διεργασίας
(Αλγόριθμος χρονοδρομολόγησης)
- Αλλαγή περιβάλλοντος λειτουργίας
(Context Switch)

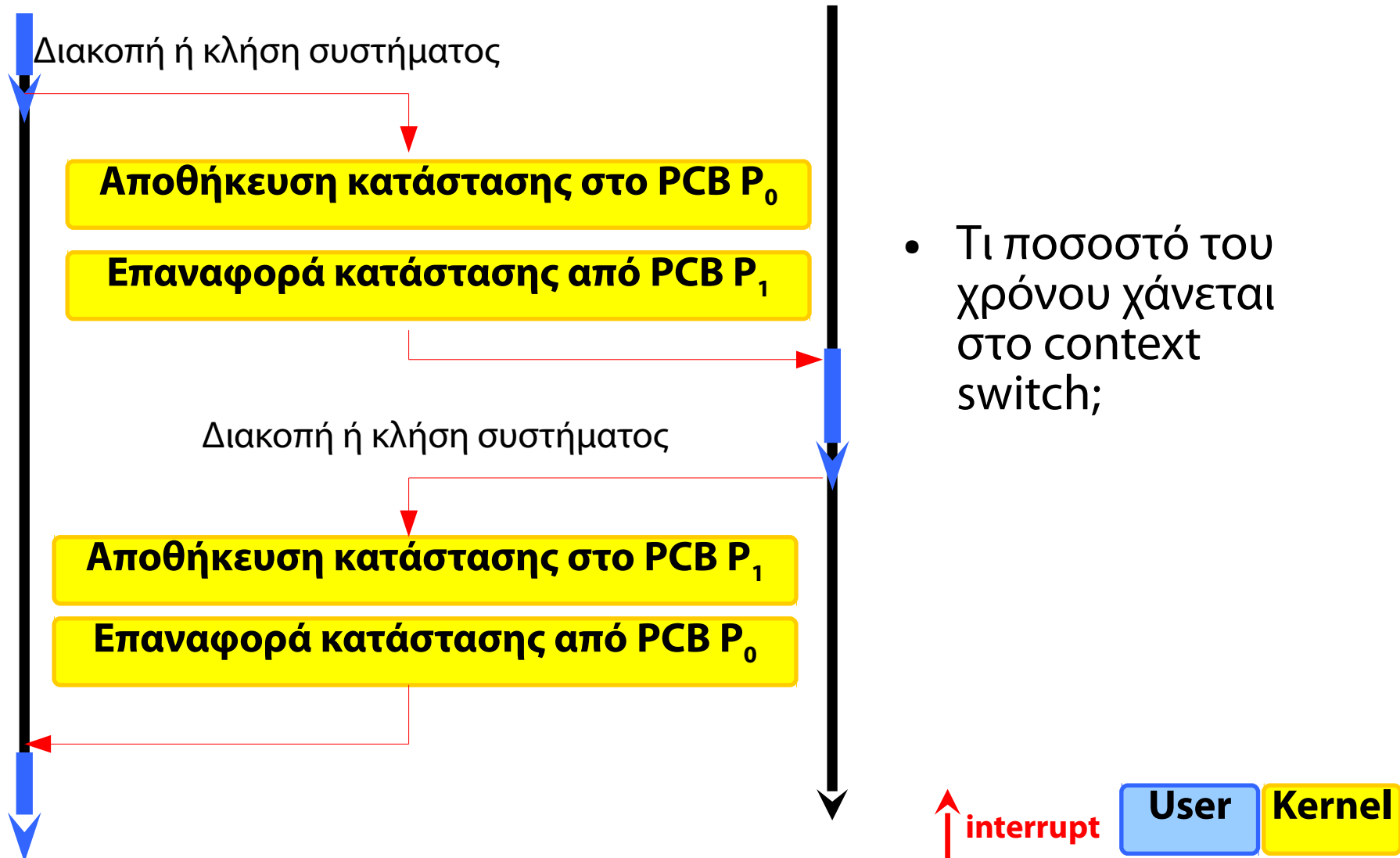
Εναλλαγή Περιβάλλοντος Λειτουργίας (1)

Context Switch

Η ΚΜΕ αλλάζει από διεργασία σε διεργασία:

- Αποθήκευση κατάστασης παλιάς (στο PCB της)
- Επαναφορά νέας διεργασίας (από το PCB της)
- Πόσο διαρκεί;
Είναι χαμένος χρόνος για τη CPU
- Πόσο συχνά;
το σύστημα πρέπει να είναι αποκρίσιμο

Εναλλαγή Περιβάλλοντος Λειτουργίας (2)

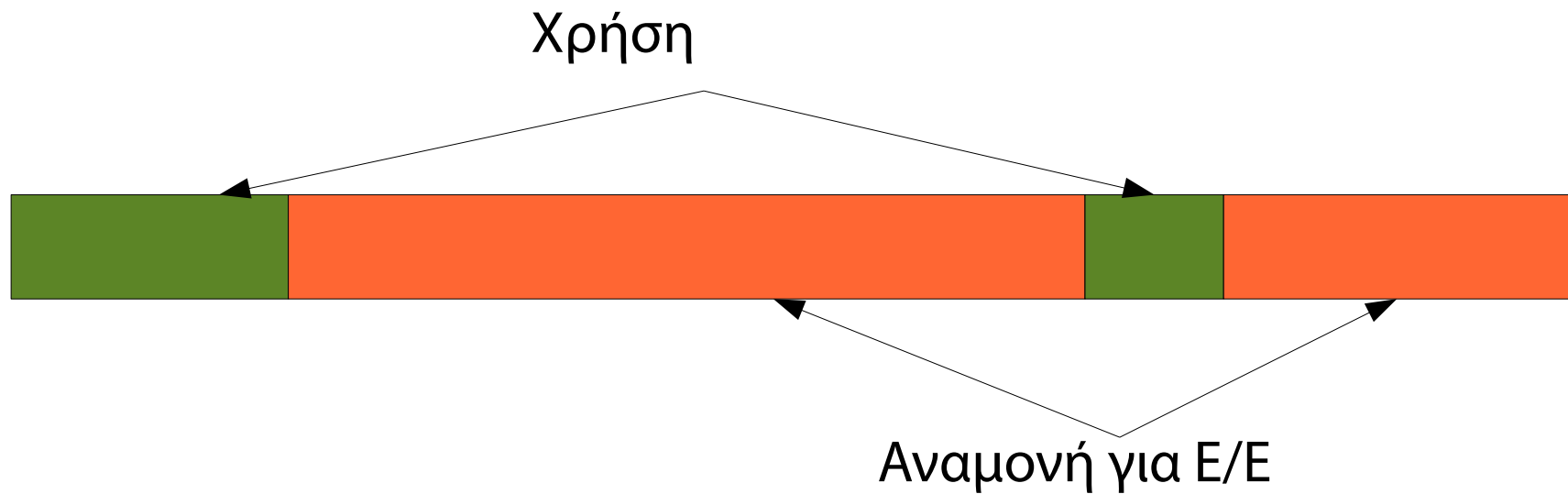


Λειτουργίες Διεργασιών

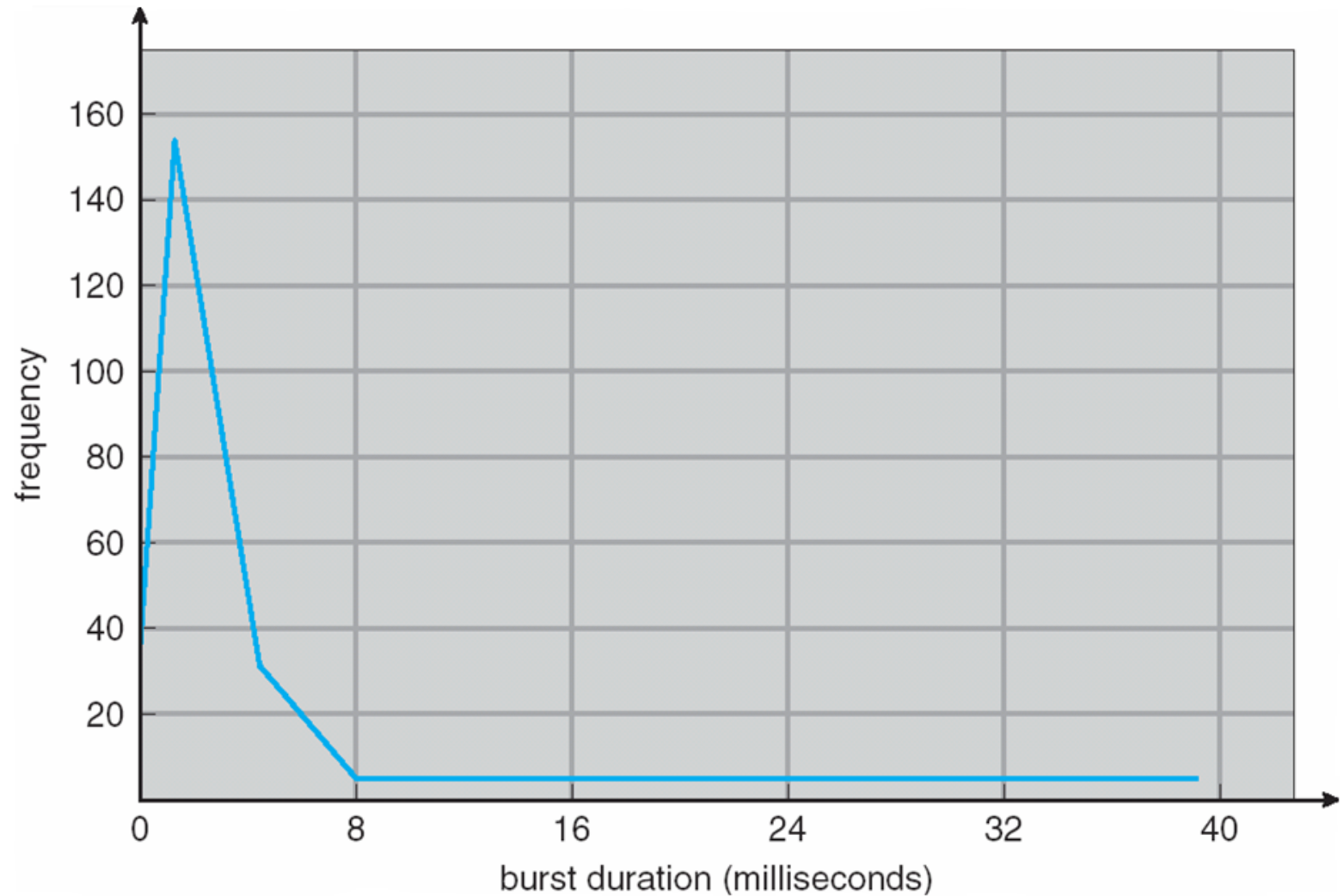
- Χρήση επεξεργαστή για Υπολογισμούς (computation)
- Αιτήσεις E/E μέσω του ΛΣ
 - Οι χρόνοι απόκρισης των μονάδων E/E είναι τάξεις μεγέθους μεγαλύτεροι από τους αντίστοιχους χρόνους του επεξεργαστή και της μνήμης.
 - Πρόσβαση στην L1: 1-2 κύκλοι
 - Πρόσβαση στη μνήμη: 100-200 κύκλοι
 - Πρόσβαση στο δίσκο: > 10.000.000 κύκλοι
 - Στα σύγχρονα συστήματα οι διαδικασίες E/E μπορούν να διεκπεραιωθούν χωρίς να απασχολείται ο επεξεργαστής. (πχ μέσω μηχανών DMA)

Συμπεριφορά Διεργασιών

- Ξεσπάσματα ΚΜΕ
- Ξεσπάσματα Ε/Ε

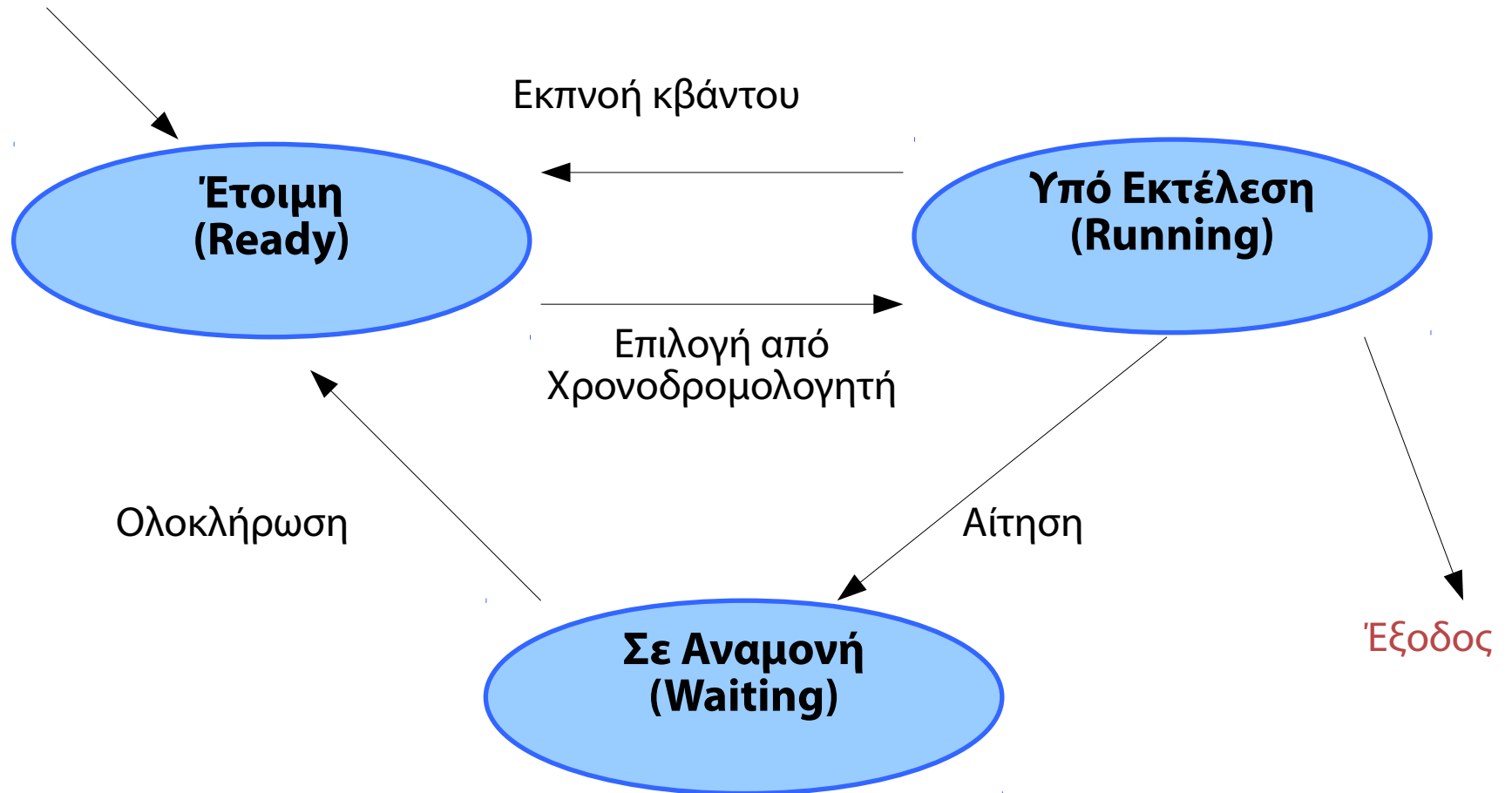


Ιστόγραμμα ξεσπασμάτων ΚΜΕ

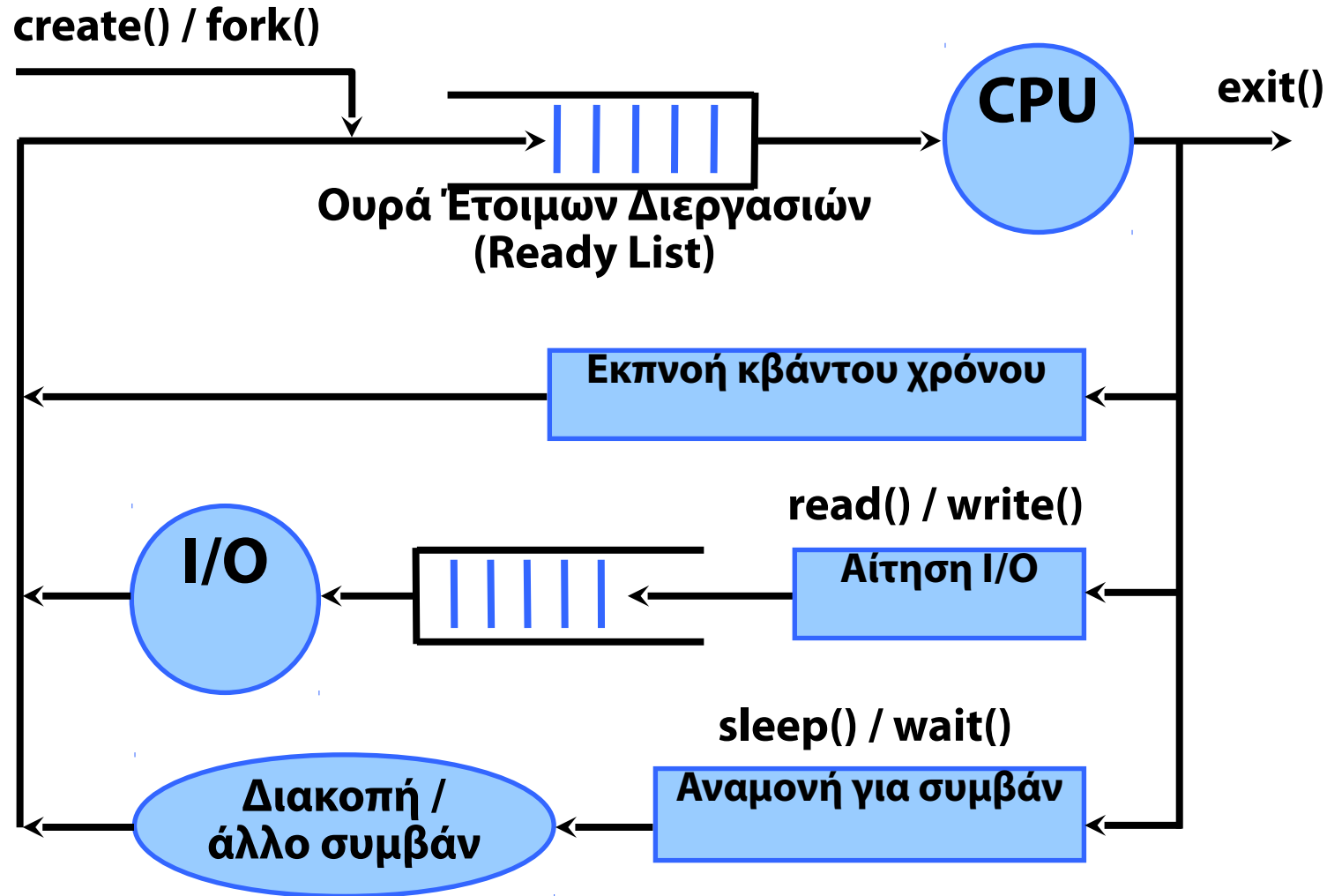


Καταστάσεις Διεργασίας

Δημιουργία



Κύκλος Ζωής μιας Διεργασίας



Χρονοδρομολογητής

Ρόλος

- Επιλογή διεργασίας (πολιτική)
- Αλλαγή περιβάλλοντος λειτουργίας – dispatching (μηχανισμός)

Στόχος:

Ανάθεση της (των) ΚΜΕ στις διεργασίες ώστε:

- Να μην χάνεται χρόνος στην ΚΜΕ εν αναμονή ολοκλήρωσης λειτουργιών E/E
- Διαμοιρασμός της ΚΜΕ στις διεργασίες, ώστε η απόκριση του συστήματος να είναι η “επιθυμητή”

Πότε εκτελείται ο χρονοδρομολογητής;

- Χωρίς διακοπές (non-preemptive) ή συνεργατικός (cooperative)
 - Όταν μια διεργασία μεταβεί σε κατάσταση αναμονής
 - Όταν μια διεργασία τερματίσει
- Διακοπτός (preemptive)
 - Όταν μια διεργασία μεταβεί σε κατάσταση εκτέλεσης (πχ ο ΧΔ ελέγχει αν η νέα διεργασία πρέπει να αντικαταστήσει την τρέχουσα)
 - Όταν πραγματοποιηθεί μια διακοπή χρονιστή

Αλλαγή διεργασίας (Dispatcher)

- Δίνει τη ΚΜΕ σε άλλη διεργασία:
 - Εναλλαγή περιβάλλοντος λειτουργίας (πχ αποκατάσταση τιμών καταχωρητών)
 - Εναλλαγή σε τρόπο λειτουργίας χρήστη (user-mode)
 - Μετάβαση στη σωστή θέση του προγράμματος (program counter register)
- Καθυστέρηση αλλαγής (dispatch latency)

Κριτήρια Αξιολόγησης

- Χρησιμοποίηση Επεξεργαστή (cpu utilization)
ποσοστό χρόνου που είναι απασχολημένη η ΚΜΕ
- Ρυθμός Διεκπεραίωσης (throughput)
διεργασίες που ολοκληρώνονται στη μονάδα χρόνου
- Χρόνος Ολοκλήρωσης (turnaround time)
χρόνος ολοκλήρωσης διεργασίας
- Χρόνος Αναμονής (waiting time)
χρόνος που η διεργασία βρίσκεται σε κατάσταση αναμονής
- Χρόνος Απόκρισης (response time)
χρόνος από την υποβολή ενός αιτήματος μέχρι να παραχθεί η πρώτη απόκριση

Επιλογή Κριτηρίων Χρονοδρομολόγησης

Εξαρτάται από τις απαιτήσεις του συστήματος

- Γενικά:
 - Μεγιστοποίηση:
 - Χρησιμοποίησης ΚΜΕ
 - Ρυθμού διεκπεραίωσης
 - Ελαχιστοποίηση Χρόνου
 - Ολοκλήρωσης
 - Αναμονής
 - Απόκρισης
- Ελάχιστες, μέγιστες, μέσες τιμές, variance
- Απαιτούνται **συμβιβασμοί!**

Παράδειγματα Κριτήριων ΧΔ

- Οικιακή χρήση (Desktop)
 - Ο χρήστης θέλει να έχει την αίσθηση ότι οι εφαρμογές του τρέχουν ταυτόχρονα
 - Το σύστημα θα πρέπει να είναι αποκρίσιμο στις πράξεις του χρήστη (π.χ. πληκτρολόγηση)
 - Γενικά απαιτείται χαμηλός χρόνος απόκρισης
- Κόμβος Υπολογισμών (Worker Node)
 - Μεγιστοποίηση ρυθμού διεκπεραίωσης
 - Ελαχιστοποίηση χρόνου ολοκλήρωσης

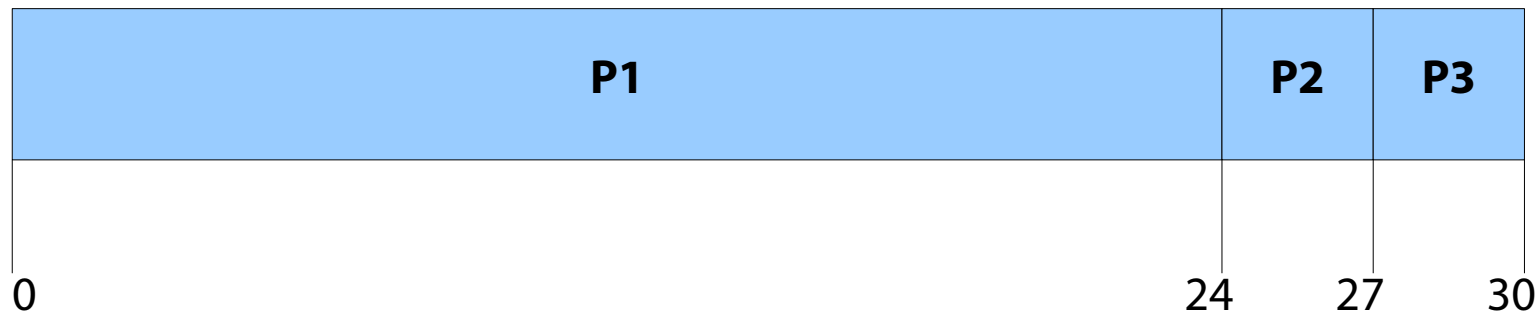
Με βάση τη σειρά άφιξης (First Come First Served – FCFS)

Χρόνοι Ξεσπάσματος ΚΜΕ

P1	24
P2	3
P3	3

Σειρά Άφιξης:

P1, P2, P3



Μέσος Χρόνος Αναμονής: $(0+24+27)/3 = 17$

Φαινόμενο Φάλαγγας (*convoy*):
όταν μικρές διεργασίες είναι πίσω από μεγάλες

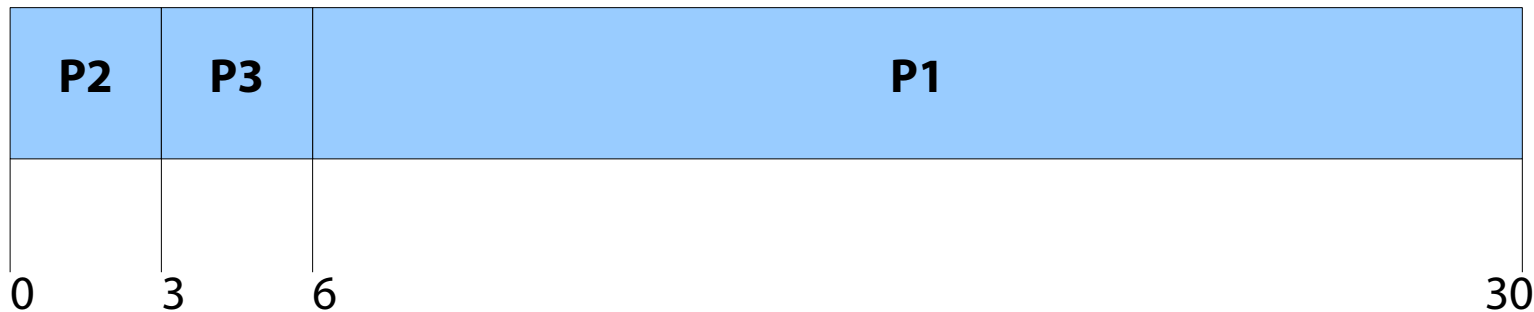
Με βάση τη σειρά άφιξης (First Come First Served – FCFS)

Χρόνοι Ξεσπάσματος ΚΜΕ

Σειρά Άφιξης:

P2, P3, P1

P1	24
P2	3
P3	3



Μέσος Χρόνος Αναμονής: $(6+0+3)/3 = 3$

Καλύτερη από προηγούμενη περίπτωση λόγω σειράς άφιξης

Με βάση τη μικρότερη Διάρκεια Εκτέλεσης (Shortest Job First - SJF)

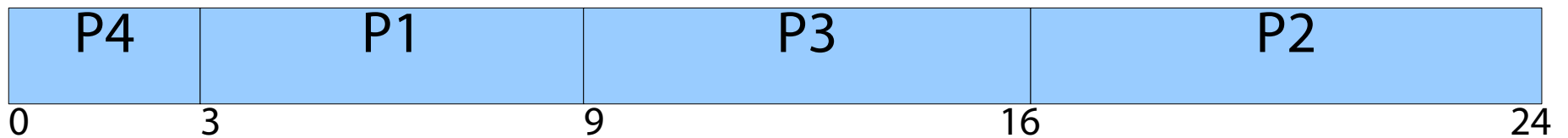
- Επιλέγεται η διεργασία με το **μικρότερο** χρόνο ξεσπάσματος ΚΜΕ.
- Βέλτιστος: επιτυγχάνει **ελάχιστο** μέσο χρόνο αναμονής

Ωστόσο:

- Το μήκος του επόμενου ξεσπάσματος ΚΜΕ δεν είναι (γενικά) γνωστό
- Προσέγγιση με βάση προηγούμενες τιμές

Παράδειγμα SJF

Διεργασία	Χρόνος Ξεπάσματος
P1	6
P2	8
P3	7
P4	3



Χρονοδρομολόγηση: P4, P1, P3, P2

Μέσος χρόνος αναμονής: $(0 + 3 + 9 + 16) / 4 = 7$

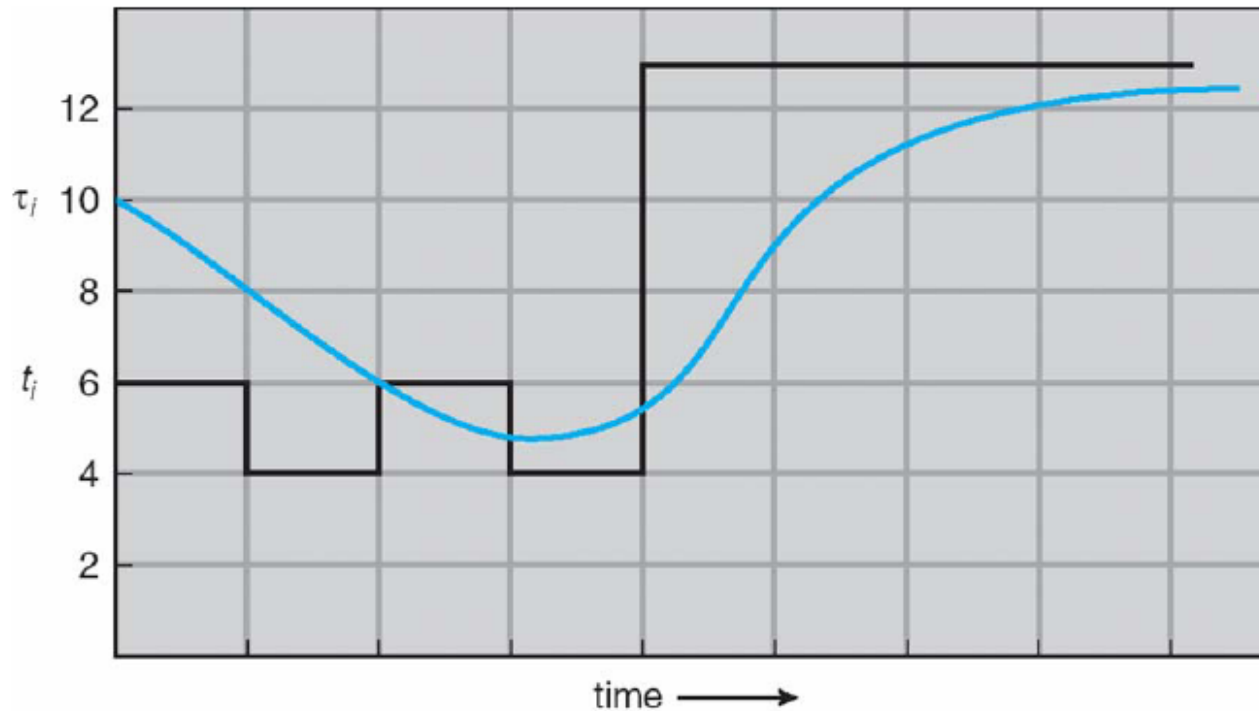
SJF: Προσέγγιση διάρκειας επόμενου ξεσπάσματος

- Γενικά δεν είναι γνωστή η διάρκεια του επόμενου ξεσπάσματος κάθε διεργασίας
- Προσέγγιση: **Εκθετικός μέσος όρος:**
- $\tau_{n+1} = \alpha t_n + (1-\alpha)\tau_n$
 - τ : πρόβλεψη
 - t : πραγματική τιμή
 - α στο $[0,1]$: σχετικό βάρος t και τ
 - $\alpha=0$: Οι μετρήσεις πραγματικών τιμών καθυστέρησης δεν επηρεάζουν
 - $\alpha=1$: Μόνο η τελευταία πραγματική καθυστέρηση μετράει

Παράδειγμα Προσέγγισης διάρκειας επόμενου ξεσπάσματος

$$a = 1/2$$

$$\tau_0 = 10$$



CPU burst (t_i)	6	4	6	4	13	13	13	...	
"guess" (τ_i)	10	8	6	6	5	9	11	12	...

SJF: διακοπτός/μη-διακοπτός

- Μη-διακοπτός

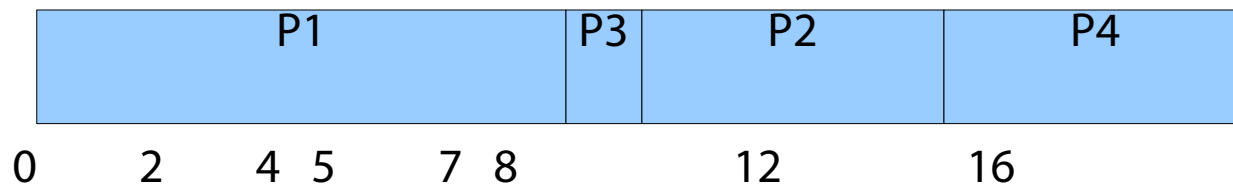
Όταν παραχωρηθεί η ΚΜΕ σε μία διεργασία, η διεργασία αυτή θα πρέπει να ολοκληρώσει τη χρήση της ΚΜΕ, έως ότου παρθεί νέα απόφαση χρονοδρομολόγησης

- Διακοπτός

Αν εισέλθει στην ουρά έτοιμων διεργασιών μια νέα διεργασία με μικρότερο αναμενόμενο χρόνο εκτέλεσης από τον εναπομείναντα της τρέχουσας, η τρέχουσα θα αντικατασταθεί (Shortest Remaining Time First - SRTF)

SJF Παράδειγμα (χωρίς διακοπές)

Διεργασία	Άφιξη	Διάρκεια
P1	0	7
P2	2	4
P3	4	1
P4	5	4



Μέσος χρόνος αναμονής: $(0 + (8-2) + (7-4) + (12-5)) / 4 = 4$

SJF Παράδειγμα (με διακοπές)

Διεργασία	Άφιξη	Διάρκεια
P1	0	7
P2	2	4
P3	4	1
P4	5	4

t	Γεγονός	Διεργασίες	Επιλογή
0-2	Άφιξη P1	(P1,7)	P1
2-4	Άφιξη P2	(P1,5) (P2,4)	P2
4-5	Άφιξη P3	(P1,5) (P2,2) (P3,1)	P3
5-7	Άφιξη P4 / Ολοκλήρωση P3	(P1,5) (P2,2) (P4,4)	P2
7-11	Ολοκλήρωση P2	(P1,5) (P4,4)	P4
11-16	Ολοκλήρωση P4	(P1,5)	P1

Μέσος Χρόνος αναμονής: $((11-2) + (5-4) + (4-4) + (7-5))/4 = 3$

Χρονοδρομολόγηση με Προτεραιότητες (priority scheduling)

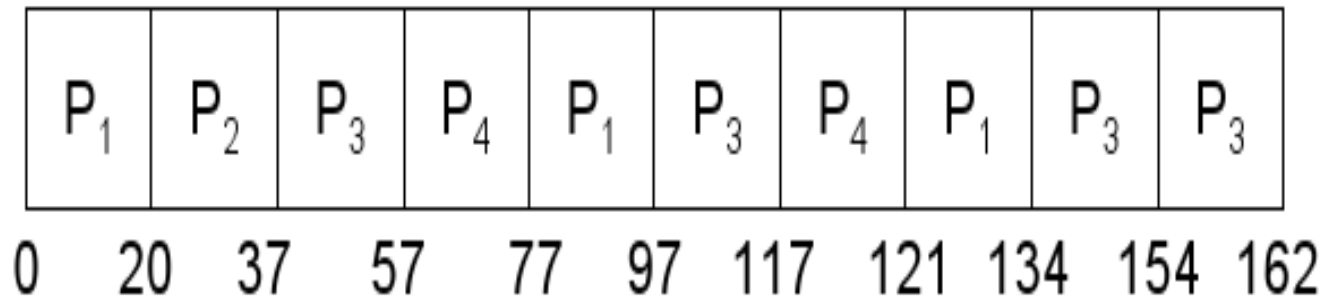
- Σε κάθε διεργασία αντιστοιχίζεται μια προτεραιότητα
- Σε κάθε απόφαση ΧΔ, επιλέγεται η διεργασία με την υψηλότερη προτεραιότητα
 - Διακοπή ΧΔ
 - Μη-διακοπή ΧΔ
- SJF Ειδική περίπτωση ΧΔ με προτεραιότητες
- Πρόβλημα **Λιμοκτονίας**
Διεργασίες με χαμηλή προτεραιότητα μπορεί να μην εκτελεστούν ποτέ
- Λύση μέσω **Γήρανσης**
Με την πάροδο του χρόνου μεγαλώνει η προτεραιότητα των διεργασιών που δεν εκτελούνται

Χρονοδρομολόγηση εκ Περιτροπής (Round Robin - RR)

- Ο ΧΔ διατηρεί μια ουρά (FIFO) για τις διεργασίες
- Διεργασίες λαμβάνουν ένα μικρό κλάσμα του χρόνου της ΚΜΕ (κβάντο χρόνου ΚΧ – time quantum)
- Με το πέρας του ΚΧ η διεργασία διακόπτεται και τοποθετείται στο τέλος της ουράς
- Επιλέγεται η διεργασία που βρίσκεται στην αρχή της ουράς

RR: Παράδειγμα

Διεργασία	Ξέσπασμα ΚΜΕ
P1	53
P2	17
P3	68
P4	24

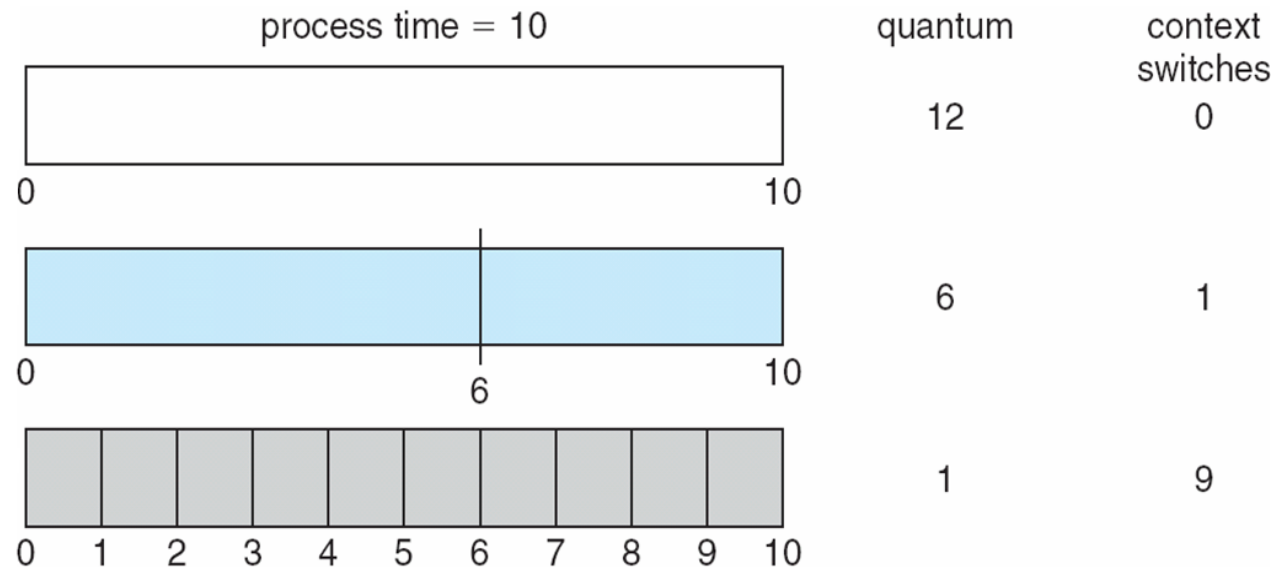


Γενικά, μεγαλύτερος χρόνος ολοκλήρωσης από τον SJF, αλλά καλύτερος χρόνος απόκρισης

RR – Κβάντο Χρόνου

- Το Κβάντο χρόνου (ΚΧ) γενικά κυμαίνεται από 10 έως 100 ms
- Αν υπάρχουν N διεργασίες στην ουρά και q το ΚΧ, καμία διεργασία δεν περιμένει παραπάνω από $(N-1)q$ χρόνο.
- Κβάντο Χρόνου
 - Μεγάλο : FCFS
 - Μικρό: Το σύστημα αναλώνεται σε context switches

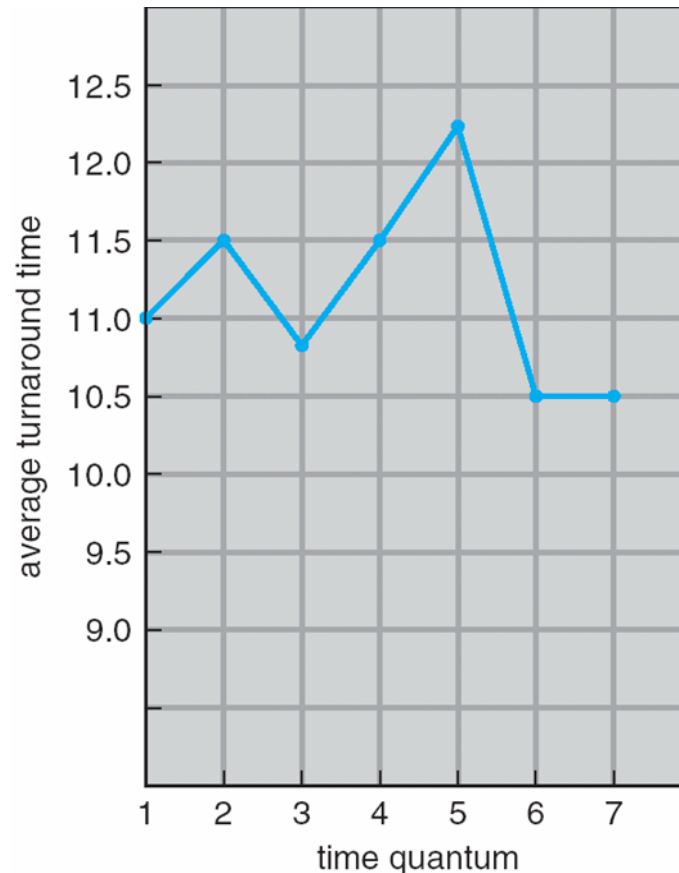
RR – ΚΧ και χρόνος εναλλαγής περιβάλλοντος λειτουργίας



Το ΚΧ πρέπει να είναι (σημαντικά) μεγαλύτερο από τον χρόνο εναλλαγής λειτουργίας

- Γενικά:
 - Χρόνος εναλλαγής $< 10\mu\text{s}$
 - ΚΧ: 10 έως 100 ms

RR – ΚΧ και χρόνος ολοκλήρωσης



process	time
P_1	6
P_2	3
P_3	1
P_4	7

- Ο χρόνος ολοκλήρωσης εξαρτάται από το ΚΧ
- Δεν είναι απαραίτητο ότι για μεγάλο ΚΧ, αυξάνεται ο χρόνος ολοκλήρωσης

Χρονοδρομολόγηση Πολυεπίπεδων Ουρών

Διαφορετικές κλάσεις διεργασιών

(ως προς τη χρονοδρομολόγηση):

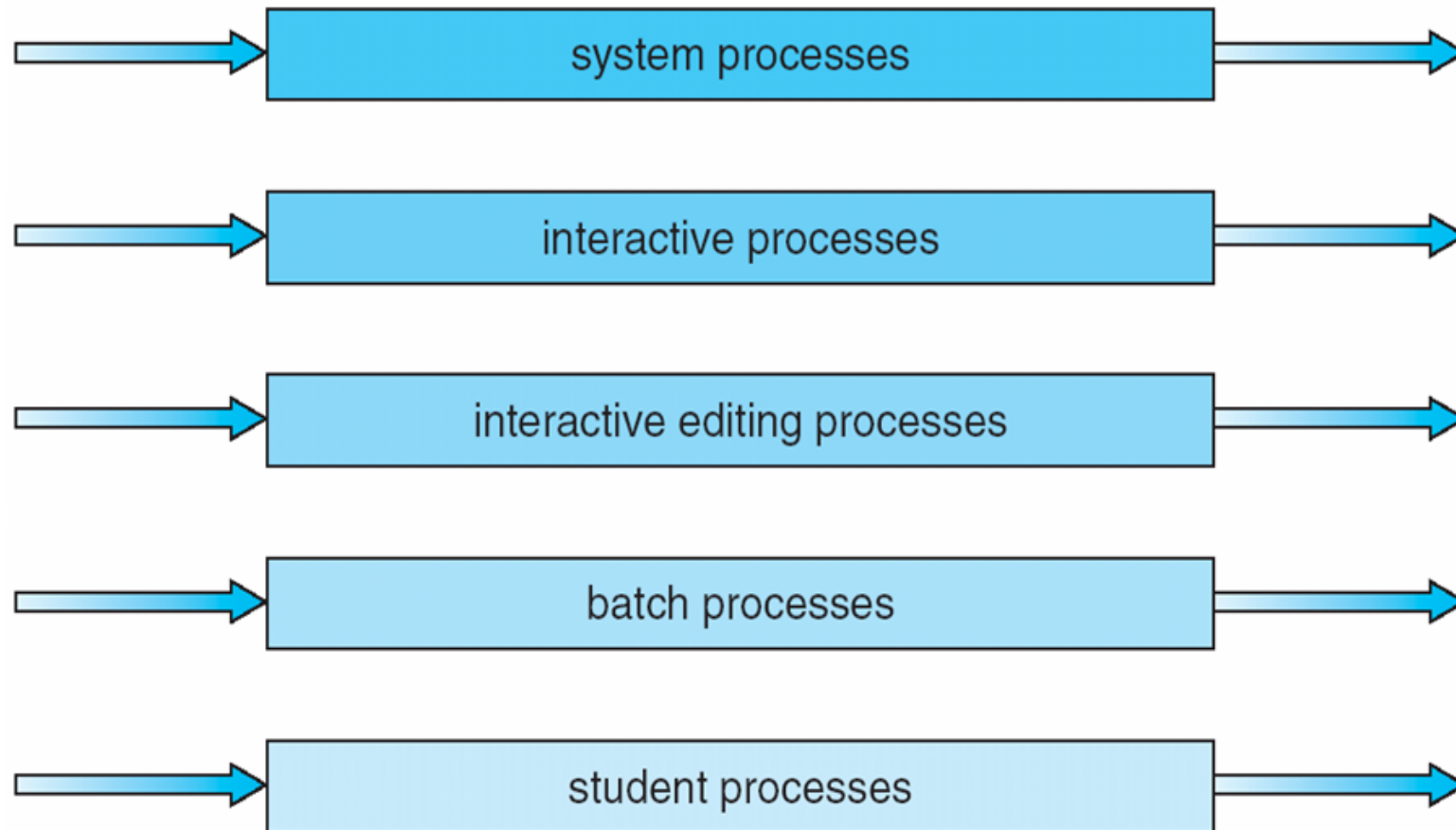
- Διεργασίες συστήματος
(υψηλή προτεραιότητα)
- Διεργασίες μαζικής επεξεργασίας (batch)
- Αλληλεπιδραστικές διεργασίες (interactive)
(σημαντικές για την εμπειρία του χρήστη)
- Διεργασίες χαμηλής προτεραιότητας
(μόνο όταν το σύστημα είναι άδειο πχ SETI@home)

Χρονοδρομολόγηση Πολυεπίπεδων Ουρών

- Η ουρά με τις έτοιμες διεργασίες χωρίζεται σε πολλαπλές ουρές, ανάλογα με την κλάση
- Κάθε ουρά έχει τον δικό της αλγόριθμο ΧΔ
- Παράδειγμα
 - Ουρά προσκηνίου (interactive), RR
 - Ουρά παρασκηνίου (batch), FCFS
- ΧΔ μεταξύ των ουρών
 - Προτεραιότητες (πιθανότητα λιμοκτονίας)
 - Κάθε ουρά λαμβάνει ποσοστό χρήσης της ΚΜΕ

Χρονοδρομολόγηση Πολυεπίπεδων Ουρών

highest priority



lowest priority

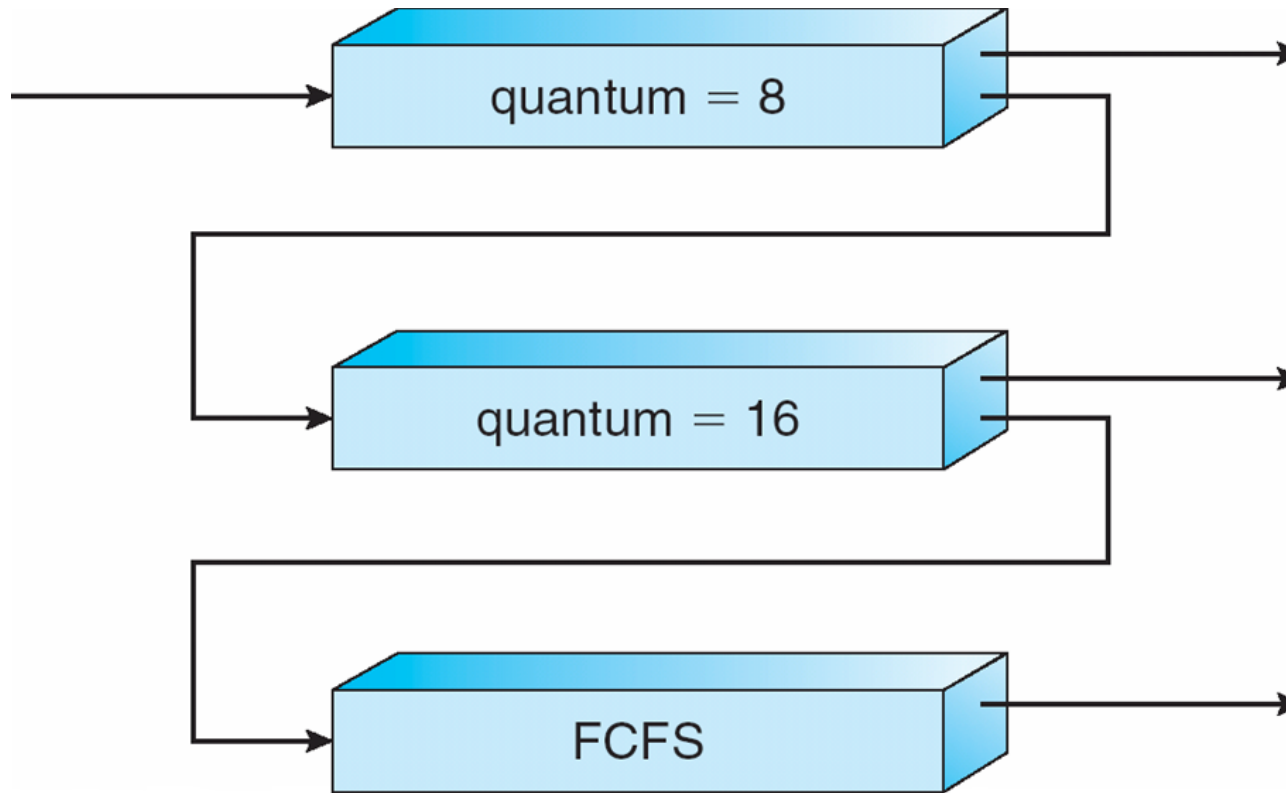
Πολυεπίπεδες Ουρές με Ανατροφοδότηση

- Ανατροφοδότηση: μετακίνηση διεργασιών μεταξύ διαφορετικών ουρών
(μπορεί πχ να χρησιμοποιηθεί για υλοποίηση γήρανσης)
- ΧΔ Πολυεπίπεδων Ουρών με Ανατροφοδότηση
 - Πλήθος ουρών
 - Τύπος ΧΔ για κάθε ουρά
 - Μέθοδοι αναβάθμισης/υποβάθμισης διεργασιών
 - Μέθοδος επιλογής ουράς για νέες διεργασίες

Πολυεπίπεδες Ουρές με Ανατροφοδότηση: Παράδειγμα

- 3 Ουρές:
 - Q0: RR με $KX=8ms$
 - Q1: RR με $KX=16ms$
 - Q2: FCFS
- Χρονοδρομολόγηση:
 - Νέα διεργασία εισέρχεται στην Q0
 - Αν δεν ολοκληρωθεί σε $8ms$ μεταφέρεται στην Q1
 - Αν δεν ολοκληρωθεί σε $16ms$ ($8+16=24$) μεταφέρεται στην Q2

Πολυεπίπεδες Ουρές με Ανατροφοδότηση: Παράδειγμα



Συστήματα Πολλαπλών Επεξεργαστών / Νημάτων

- Επεξεργαστές
 - Πολλαπλοί πυρήνες
 - Φυσικές ΚΜΕ
 - Μοιράζονται μέρος ιεραρχίας μνήμης
 - Πολλαπλά νήματα σε έναν πυρήνα
 - Λογικές ΚΜΕ
 - Μοιράζονται μονάδες (πχ ALU)
- Συστήματα πολλαπλών επεξεργαστών μοιραζόμενης μνήμης
 - Ίδια ταχύτητα πρόσβασης στη μνήμη για κάθε ΚΜΕ
 - Διαφορετική ταχύτητα πρόσβασης στη μνήμη για κάθε ΚΜΕ (NUMA)

Χρονοδρομολόγηση σε Συστήματα Πολλαπλών Επεξεργαστών

- Ασύμμετρη
 - Ο ΧΔ τρέχει σε έναν επεξεργαστή
 - Οι υπόλοιποι χρησιμοποιούνται για εκτέλεση κώδικα χρήστη
- Συμμετρική
 - Ο ΧΔ τρέχει σε όλους τους επεξεργαστές
 - Κάθε ΧΔ επιλέγει διεργασία προς εκτέλεση στον αντίστοιχο επεξεργαστή
 - Απαιτείται συγχρονισμός

Ζητήματα σε ΧΔ σε Συστήματα Πολλαπλών Επεξεργαστών

- Προσκόλληση σε Επεξεργαστή
(processor affinity)
 - Επίδοση κρυφής μνήμης
- Εξισορρόπηση Φόρτου
(load balancing)
 - Απαιτεί μετακίνηση διεργασιών
 - Μετακίνηση ώθησης (push migration)
 - Μετακίνηση έλξης (pull migration)
- Γνώση φυσικής τοπολογίας συστήματος
 - Πχ NUMA, SMT

Χρονοδρομολόγηση Νημάτων

- Νήματα χώρου πυρήνα
 - Χρονοδρομολογούνται από ΛΣ
 - Υψηλό (σχετικά) κόστος διαχείρισης
 - Απλούστερη Υλοποίηση
- Νήματα χώρου χρήστη
 - Χρονοδρομολόγηση προσαρμοσμένη στην εφαρμογή
 - Χαμηλό (σχετικά) κόστος διαχείρισης
 - Διακοπή / συνεργατική χρονοδρομολόγηση
 - Ζητήματα:
 - Τι γίνεται με κλήσεις συστήματος που προκαλούν αναστολή εκτέλεσης (πχ read())
 - Εκμετάλλευση πολλαπλών επεξεργαστών

Συστήματα Πραγματικού Χρόνου (Real-Time systems)

- Αυστηρά (hard real-time systems)
 - Πρέπει να είναι σε θέση να ολοκληρώσουν μια διεργασία σε αυστηρά χρονικά περιθώρια
 - Χρησιμοποιούν δέσμευση πόρων
 - Δύσκολο να υλοποιηθούν
 - Συστήματα ειδικού σκοπού (Ιατρικές εφαρμογές)
- Χαλαρά (soft real-time systems)
 - Δίνουν προτεραιότητα στις κρίσιμες (real-time) διεργασίες, έναντι των υπολοίπων
 - Πχ εφαρμογές πολυμέσων σε συστήματα γενικού σκοπού