ΑΣΚΗΣΗ “ΠΑΡΑΛΛΑΓΗΣ” M/M/1

- Η άσκηση αυτή σκοπό έχει να μας δείξει ότι με βασικές γνώσεις επί των συστημάτων εξυπηρέτησης (π.χ. ουρά αναμονής M/M/1) μπορούμε να αντιμετωπίσουμε όχι μόνο τυποποιημένα θέματα (επιλυόμενα με άμεση εφαρμογή της θεωρίας) αλλά και ιδιότροπα θέματα (συστήματα εξυπηρέτησης).
- Θεωρήστε μια παραλλαγή της ουράς αναμονής M/M/1 όπου ο εξυπηρετητής (server) αρχίζει να εξυπηρετεί τις κλήσεις (πελάτες) από την στιγμή που ο αριθμός των πελατών στο σύστημα θα γίνει 3. Διαφορετικά, το σύστημα συμπεριφέρεται όπως στην ουρά M/M/1 με μέσο ρυθμό άφιξης των πελατών λ και μέσο ρυθμό εξυπηρέτησης μ. Σημειωτέον ότι ο εξυπηρετητής αφ’ ης στιγμής αρχίσει να εξυπηρετεί τις κλήσεις, τις εξυπηρετεί κανονικά, ακόμη και όταν ο αριθμός τους είναι μικρότερος του 3. Αν όμως εξυπηρετηθούν όλες οι κλήσεις και το σύστημα ευρεθεί στην κατάσταση $n = 0$ (κενό), τότε πάλι ο εξυπηρετητής αρχίζει να εξυπηρετεί τις κλήσεις από την στιγμή που ο αριθμός τους στο σύστημα θα γίνει 3.

ΖΗΤΕΙΤΑΙ:

α) Να σχεδιάσετε το διάγραμμα καταστάσεων του συστήματος.

β) Χρησιμοποιώντας την έννοια της σφαιρικής ισορροπίας (global balance), σύμφωνα με την οποία ο (συνολικός) ρυθμός εξόδου από μια κατάσταση *n* ισούται με τον (συνολικό) ρυθμό εισόδου στην κατάσταση αυτή (όταν το σύστημα βρίσκεται σε κατάσταση ισορροπίας), να γράψετε τις εξισώσεις σφαιρικής ισορροπίας για το σύστημα αυτό με σκοπό να υπολογίσετε, τις πιθανότητες να υπάρχουν στο σύστημα *n* πελάτες, με $n = 0, 1, 2, 3, \ldots$, συναρτήσει του φορτίου κίνησης $a = \lambda/\mu$.

γ) Να υπολογιστεί η πιθανότητα να είναι κατειλημμένος ο εξυπηρετητής. Σχολιάστε το αποτέλεσμα, σε σχέση με την αντίστοιχη πιθανότητα του κλασσικού συστήματος M/M/1;

δ) Έστω ότι το προσφερόμενο φορτίο κίνησης στο σύστημα είναι $a = 0.5$ erl. Να ευρεθεί η πιθανότητα στο σύστημα να έχουμε 3 κλήσεις (πελάτες), δηλ. $n = 3$, καθώς και η πιθανότητα ο εξυπηρετητής να είναι κατειλημμένος.

ΥΠΟΔΕΙΞΗ: Εκφράσετε τις πιθανότητες των καταστάσεων μέχρι και την κατάσταση n=3 συναρτήσει της κατάστασης n=0. Τις πιθανότητες των καταστάσεων n, για n>3, θα τις εκφράσετε συναρτήσει της κατάστασης n=3 (θα έχετε βρει ότι η n=3 εκφράζεται συναρτήσει της κατάστασης n=0). Για να βρείτε την πιθανότητα της κατάστασης n=0 θα πάρετε την συνθήκη κανονικοποίησης των πιθανοτήτων, δηλ. το άθροισμά τους (όλων των πιθανοτήτων, μέχρι το άπειρον) είναι 1 (μονάδα). Σημειωτέον, ο όρος $\sum_{n=0}^{\infty} a^n$ (άθροισμα γεωμετρικής προόδου) συγκλίνει στην τιμή 1 / (1 - α), όταν α < 1.

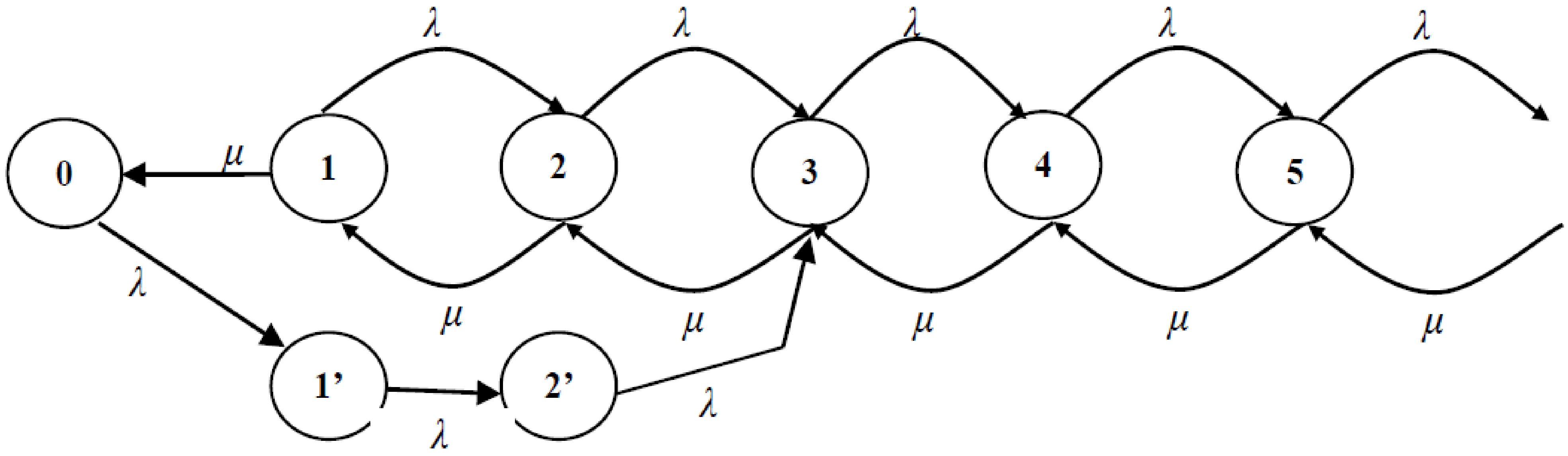
ΛΥΣΗ

α) Αρχίζουμε (πάντα) από την κατάσταση $n = 0$ (όπου το σύστημα είναι άδειο). Ακολούθως, με άφιξη μιας κλήσης (με ρυθμό λ κλήσεις/s) μεταβαίνουμε από την κατάσταση 0 στην κατάσταση 1’, μετά, με συνακόλουθη άφιξη δεύτερης κλήσης, στην κατάσταση 2’ και μετά με μία ακόμη άφιξη κλήσης στην κατάσταση 3 (βλέπε σχήμα).

Σύμφωνα με την «εκφώνηση» από τις καταστάσεις 1’ και 2’ δεν υπάρχει μετάβαση προς τα πίσω (από 1’ → 0 ή από 2’ → 1’), διότι δεν υπάρχει εξυπηρέτηση από τον εξυπηρετητή ακόμη. Ο εξυπηρετητής εξυπηρετεί τις κλήσεις (με ρυθμό μ κλήσεις/s) από την κατάσταση $n = 3$.

Έτσι από την κατάσταση $n = 3$ με άφιξη κλήση μεταβαίνουμε στην κατάσταση $n = 4$, ενώ με εξυπηρέτηση μιας κλήσης θα βρεθούμε στην κατάσταση $n = 2$, απ’ όπου μπορούμε να κινηθούμε πλέον προς τα εμπρός (forward) ή προς τα πίσω (backward) όπως στο κανονικό M/M/1. Αν με εξυπηρέτηση μιας κλήσης από 2 → 1 και ομοίως από 1 → 0, τότε πρέπει πάλι (σύμφωνα με την εκφώνηση) από 0 → 1’, από 1’ → 2’ και από 2’ → 3.

Οπότε σχηματίζουμε το διάγραμμα καταστάσεων του συστήματος που φαίνεται στο σχήμα που ακολουθεί.



Σχήμα: Διάγραμμα καταστάσεων

β) Από την σφαιρική ισορροπία κάθε κατάστασης (global balance: «Συνολικός ρυθμός εισόδου σε μία κατάσταση = Συνολικό ρυθμό εξόδου από την κατάσταση αυτή») και αφού το φορτίο κίνησης α ισούται προς λ / μ , προκύπτει ότι:

- Για την κατάσταση $n = 0$: $\mu P_1 = \lambda P_0 \Rightarrow P_1 = aP_0$ (Βάσει του σχήματος, η είσοδος στην κατάσταση 0 γίνεται από την κατάσταση 1 με ρυθμό μ, ενώ η έξοδος με ρυθμό λ οδηγεί στην κατάσταση 1').
- Για την κατάσταση $n = 1'$: $\lambda P_0 = \lambda P_{1'}$
- Για την κατάσταση $n = 2'$: $\lambda P_{1'} = \lambda P_{2'}$ άρα $\lambda P_0 = \lambda P_{1'} = \lambda P_{2'} \Rightarrow P_0 = P_{1'} = P_{2'}$
- Για την κατάσταση $n = 1$: $\mu P_2 = \lambda P_1 + \mu P_1 \Rightarrow \mu P_2 = (\lambda+\mu)P_1 \Rightarrow P_2 = a(a+1)P_0$
- Για την κατάσταση $n = 2$: $\mu P_3 + \lambda P_1 = \lambda P_2 + \mu P_2 \Rightarrow \mu P_3 = -\lambda P_1 + (\lambda+\mu)P_2 \Rightarrow$

  $P_3 = -\alpha P_1 + (\alpha+1)P_2 = -\alpha^2 P_0 + (\alpha+1)(\alpha+1)\alpha P_0 \Rightarrow P_3 = (a^2+a+1)aP_0$
- Για την κατάσταση $n = 3$: $\mu P_4 + \lambda P_{2'} + \lambda P_2 = \mu P_3 + \lambda P_3 \Rightarrow P_4 + \alpha P_{2'} + \alpha P_2 = \alpha P_3 + \alpha P_3 \Rightarrow P_4 =$

  $= (a+1)P_3 - aP_0 - a(\alpha+1)\alpha P_0 = (\alpha+1)(a^2+a+1)aP_0 - \alpha P_0 - a^3 P_0$

  $- a^2 P_0 \Rightarrow P_4 = (a^4 + a^3 + a^2 + a^3 + a^2 + \alpha)P_0 - \alpha P_0 - a^3 P_0 - a^2 P_0 \Rightarrow$

  $P_4 = (a^4 + a^3 + a^2)P_0 = a^2(a^2 + \alpha + 1)P_0 \Rightarrow P_4 = \alpha P_3$

- Για την κατάσταση $n = 4$: $\mu P_5 + \lambda P_3 = \lambda P_4 + \mu P_4 \Rightarrow P_5 = (a+1)P_4 - aP_3 \Rightarrow$

  $P_5 = (a+1)\alpha P_3 - aP_3 = \alpha P_3(\alpha + 1 - 1) \Rightarrow P_5 = a^2 P_3$
- Για την κατάσταση $n = 5$: $\mu P_6 + \lambda P_4 = \lambda P_5 + \mu P_5 \Rightarrow P_6 = (a+1)P_5 - aP_4 \Rightarrow$

  $P_6 = (a+1)a^2 P_3 - a^2 P_3 = a^2 P_3(\alpha + 1 - 1) \Rightarrow P_6 = a^3 P_3$

  …
- Για την κατάσταση n γενικά: $P_n = a^{n-3} P_3$ με $n > 3$ (δηλ. $n = 4, 5, 6, \dots$)

Γνωρίζουμε επίσης ότι το άθροισμα όλων των πιθανοτήτων ισούται με 1 (συνθήκη κανονικοποίησης των πιθανοτήτων). Αν την εφαρμόσουμε στο σύστημά μας θα δούμε ότι μπορούμε να υπολογίσουμε την πιθανότητα P_0 το σύστημα να είναι άδειο:

$$\sum_{n=0}^{\infty} P_n = 1 \Rightarrow P_0 + P_{1'} + P_{2'} + P_1 + P_2 + P_3 + P_4 + P_5 + \ \ldots = 1 \Rightarrow$$

$$P_0 + P_0 + P_0 + \alpha P_0 + a(a+1)P_0 + \ a(a^2+a+1)P_0 + a^2(a^2+a+1) + \ a^3(a^2+a+1) + \ldots = 1 \Rightarrow$$

- $P_0(3 + a + a(a+1) + a(a^2+a+1) + a^2(a^2+a+1) + a^3(a^2+a+1) + \ldots) = 1 \Rightarrow$

$$P_0\left(3 + a + a(a+1) + \sum_{n=3}^{\infty} a^{n-2}(a^2+a+1)\right) = 1 \Rightarrow$$

$$P_0\left(3 + a + a(a+1) + \frac{(a^2+a+1)}{a^2}\sum_{n=3}^{\infty} a^n\right) = 1 \Rightarrow$$

$$P_0\left(3 + a + a(a+1) + \frac{(a^2+a+1)}{a^2}\left(\frac{1}{1-a} - 1 - a - a^2\right)\right) = 1 \Rightarrow$$

$$P_0\left(3 + a + a(a+1) + \frac{(a^2+a+1)a}{1-a}\right) = 1 \Rightarrow P_0 = \frac{1-a}{3}$$

Βασιζόμενοι στις προηγούμενες σχέσεις υπολογίζουμε τις πιθανότητες $P\{n\}$ να υπάρχουν στο σύστημα n πελάτες, με $n = 1, 2, 3,\ldots$ως ακολούθως:

$$P\{1\} = P_1 + P_{1'} = (1+a)P_0$$

$$P\{2\} = P_2 + P_{2'} = (1+a+a^2)P_0$$

$$P\{3\} = P_3 = (1+a+a^2)aP_0$$

$$P\{n\} = (1+a+a^2)a^{n-2}P_0, \quad n = 2, 3, 4,\ldots$$

γ) Ο εξυπηρετητής είναι κατειλημμένος σε όλες τις καταστάσεις εκτός από τις 0, 1' και 2'. Επομένως η πιθανότητα να είναι ο εξυπηρετητής κατειλημμένος ισούται με:

$$P_{busy\ server} = 1 - P_0 - P_{1'} - P_{2'} = 1 - 3P_0 = a \ .$$

Παρά την διαφορετική συμπεριφορά του συγκεκριμένου συστήματος, ο εξυπηρετητής είναι το ίδιο πιθανό να είναι κατειλημμένος με την περίπτωση της κλασσικής ουράς M/M/1.

- Στο κλασσικό M/M/1 η πιθανότητα ο server να είναι κατειλημμένος δίδεται από το φορτίο κίνησης που διεκπεραιώνει ο server. Αλλά το φορτίο αυτό είναι ίσο με το προσφερόμενο φορτίο α αφού δεν έχουμε απώλεια κίνησης παρά μόνον καθυστέρηση. Το ίδιο ακριβώς συμβαίνει και στο σύστημα του θέματος αυτού. Επομένως το γεγονός ότι βρήκαμε με αναλυτικό τρόπο (με μαθηματικές σχέσεις) ότι στο σύστημά μας η πιθανότητα ο server να είναι κατειλημμένος είναι ίση προς α, δείχνει και την ορθότητα της ανάλυσης.

δ) $P\{3\} = P_3 = (1 + a + a^2)\, aP_0 = (1 + a + a^2)\, a(1 - a) / 3 = (1 + 0.5 + 0.25)\cdot 0.5\cdot(1 - 0.5) / 3 = 0.14583 \Rightarrow$

$\boldsymbol{P_3 = 14.58\%}$

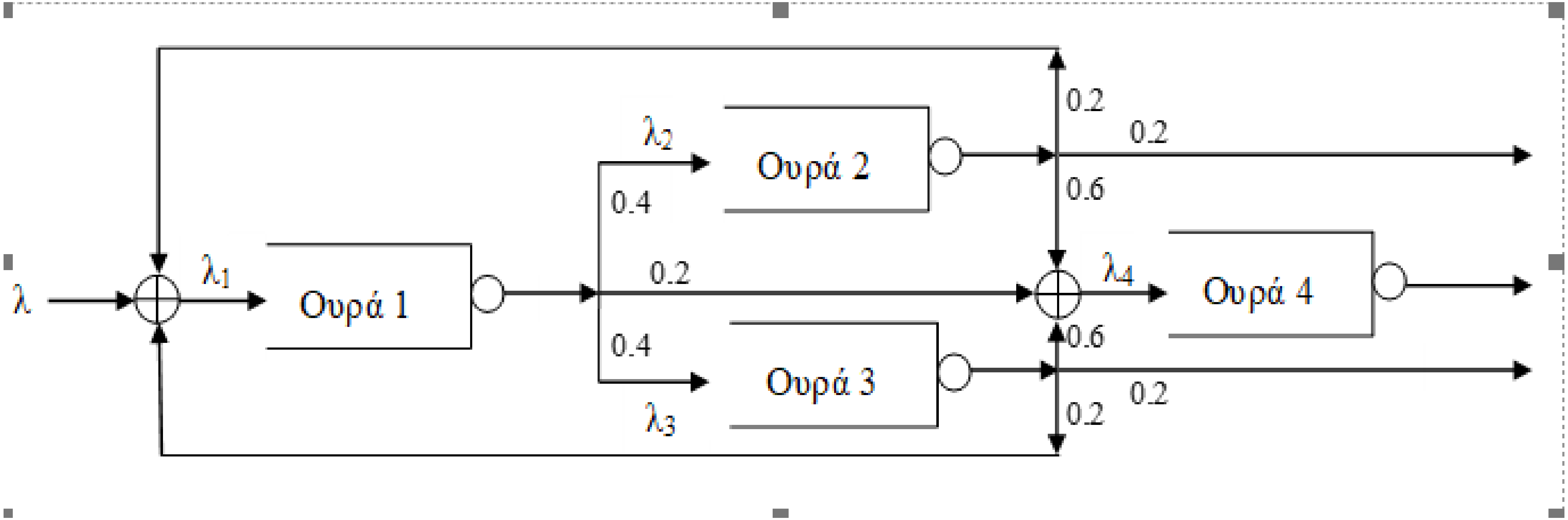
$P_{busy\ server} = a = 0.5 \Rightarrow \boldsymbol{P_{busy\ server} = 50\%}$.

ΔΙΚΤΥΟ ΑΝΑΜΟΝΗΣ

- Θεωρήστε το δίκτυο αναμονής του παρακάτω σχήματος, όπου πακέτα δεδομένων φτάνουν από εξωτερική πηγή στο υποσύστημα 1 (δηλ. Ουρά 1) με ρυθμό λ, και αναχωρούν από τα υποσυστήματα 2, 3 και 4 (δηλ. Ουρές 2, 3 και 4). Ο ρυθμός άφιξης των πακέτων στα υποσυστήματα 1, 2, 3 και 4 (Ουρές 1, 2, 3 και 4) ορίζεται ως λ_1 , λ_2 , λ_3 και λ_4 αντίστοιχα. Οι ρυθμοί εξυπηρέτησης των πακέτων για κάθε υποσύστημα-ουρά είναι $\mu_1 = \mu$ και $\mu_2 = \mu_3 = \mu_4 = 0.5\mu$. Θεωρήστε ότι όλες οι ουρές είναι M/M/1.

- A. Προσδιορίστε τον μέγιστο ρυθμό αφίξεων λ ώστε το (όλο) σύστημα να βρίσκεται σε ισορροπία (steady state).

- B. Έστω ότι $\lambda = 0.1$ πακέτα/s και $\mu = 1$ πακέτο/s. Να προσδιορίσετε το μέσο επίπεδο χρήσης (utilization = φορτίο κίνησης ανά εξυπηρετητή) κάθε υποσυστήματος-ουράς, το μέσο πλήθος πακέτων N_i σε κάθε υποσύστημα-ουρά i (i=1,…,4) αλλά και στο δίκτυο (σύστημα των τεσσάρων ουρών) συνολικά. Προσδιορίστε επίσης τον μέσο χρόνο προσπέλασης του συνολικού συστήματος από τα εισερχόμενα πακέτα.


λ
λ_1
Ουρά 1
0.4
λ_2
Ουρά 2
0.2
0.4
λ_3
Ουρά 3
0.2
0.2
0.6
λ_4
Ουρά 4
0.6
0.2
0.2

- Γ. Το εν λόγω δίκτυο αναμονής μελετάται ως ένα δίκτυο αναμονής μορφής γινομένου (product form). Αν δηλαδή ορίσουμε ως n_i το πλήθος των πακέτων στο υποσύστημα i (i=1,…,4) και $P_i(n_i)$ την αντίστοιχη πιθανότητα τότε η από κοινού πιθανότητα $P(n_1, n_2, n_3, n_4)$, (δηλ. να έχουμε συγχρόνως n_1 κλήσεις στο 1ο υποσύστημα, n_2 κλήσεις στο 2ο υποσύστημα, n_3 κλήσεις στο 3ο υποσύστημα και n_4 κλήσεις στο 4ο υποσύστημα), μπορεί να υπολογιστεί πολλαπλασιάζοντας τις αντίστοιχες πιθανότητες των επιμέρους τεσσάρων υποσυστημάτων M/M/1 (θεώρημα Jackson):

$$P(n_1, n_2, n_3, n_4) = P_1(n_1)^* P_2(n_2)^* P_3(n_3)^* P_4(n_4)$$

- όπου $P_i(n_i) = (1-\alpha_i)\alpha_i^{n_i}$ είναι η πιθανότητα να έχουμε n_i πακέτα στο υποσύστημα i (i=1,…,4), που δέχεται (και διεκπεραιώνει) φορτίο κίνησης $\alpha_i = \lambda_i / \mu_i$.
- Βασιζόμενοι στην παραπάνω σχέση, να υπολογίσετε την πιθανότητα P(0,0,0,0), δηλαδή την πιθανότητα το δίκτυο αναμονής να είναι εντελώς άδειο.
- *Υπόδειξη*: Στο αναφερόμενο δίκτυο αναμονής ισχύουν ότι α) ο ρυθμός άφιξης σε μια ουρά είναι ίσος με τον ρυθμό αναχώρησης από το υποσύστημα (θεώρημα Burke), β) σε περίπτωση πολλαπλών (Poisson) ροών προς μια ουρά, η συνολική ροή αποτελεί το άθροισμα των επιμέρους ροών, και γ) μια ροή Poisson διασπάται σε επιμέρους ροές Poisson.

ΛΥΣΗ

- **Α.** Βασιζόμενοι στο σχήμα και την υπόδειξη, προκύπτει το παρακάτω σύστημα εξισώσεων:
  - $\lambda_1 = \lambda + 0.2\lambda_2 + 0.2\lambda_3$
  - $\lambda_2 = 0.4\lambda_1$
  - $\lambda_3 = 0.4\lambda_1$
  - $\lambda_4 = 0.2\lambda_1 + 0.6\lambda_2 + 0.6\lambda_3$

ή $\lambda_1 = \lambda + 0.2*0.4\,\lambda_1 + 0.2*0.4\,\lambda_1 \Rightarrow \lambda = \lambda_1 - 0.08\,\lambda_1 - 0.08\,\lambda_1 \Rightarrow \lambda = 0.84\,\lambda_1$

$\lambda_4 = 0.2\,\lambda_1 + 0.6*0.4\,\lambda_1 + 0.6*0.4\,\lambda_1 \Rightarrow \lambda_4 = 0.2\,\lambda_1 + 0.24\,\lambda_1 + 0.24\,\lambda_1 \Rightarrow \lambda_4 = 0.68\,\lambda_1$

- Οπότε οδηγούμαστε στην επίλυση του συστήματος λαμβάνοντας τις τιμές:
  - $\lambda_1 = 1.19048\,\lambda$
  - $\lambda_2 = 0.47619\,\lambda$
  - $\lambda_3 = 0.47619\,\lambda$
  - $\lambda_4 = 0.80952\,\lambda$

- Προκειμένου το σύστημα να είναι σε ισορροπία, θα πρέπει το επίπεδο χρήσης (utilization ρ) σε κάθε μία από τις 4 ουρές να είναι μικρότερο του 1. Θέτοντας $\rho = \lambda / \mu$ και επειδή από την εκφώνηση της άσκησης ισχύει $\mu_1 = \mu$ και $\mu_2 = \mu_3 = \mu_4 = 0.5\mu$, το επίπεδο χρήσης σε κάθε ουρά είναι:
  - $\rho_1 = \lambda_1 / \mu_1 = (1.19048\ \lambda) / (\mu) = 1.19048\ (\lambda / \mu) = 1.19048\ \rho$
  - $\rho_2 = \lambda_2 / \mu_2 = (0.47619\ \lambda) / (0.5\mu) = 0.95238\ (\lambda / \mu) = 0.95238\ \rho$
  - $\rho_3 = \lambda_3 / \mu_3 = (0.47619\ \lambda) / (0.5\mu) = 0.95238\ (\lambda / \mu) = 0.95238\ \rho$
  - $\rho_4 = \lambda_4 / \mu_4 = (0.80952\ \lambda) / (0.5\mu) = 1.61904\ (\lambda / \mu) = 1.61904\ \rho$

  Έχοντας πλέον όλα επίπεδα χρήσης ρ_i (i=1,2,3,4) ανάλογα του ρ, παρατηρούμε ότι η Ουρά 4 εμφανίζει το μεγαλύτερο επίπεδο χρήσης και απαιτούμε η τιμή του ρ_4 να μην υπερβαίνει το 1. Δηλαδή:

$$\rho_4 < 1 \Rightarrow 1.61904\ \rho < 1 \Rightarrow \lambda < 0.61765\ \mu$$

- **Β.** Επειδή $\lambda = 0.1$ πακέτα/s και $\mu = 1$ πακέτο/s, έχουμε $\rho = \lambda / \mu = 0.1$. Οπότε από τις ανωτέρω τιμές των ρ_i (i=1,…,4) προκύπτει η τελική τιμή του μέσου επιπέδου χρήσης (φορτίο κίνησης ανά εξυπηρετητή) κάθε ουράς είναι**:**
  - $\boldsymbol{\rho_1 = 0.119048} = \alpha_1$ (αφού s=1)
  - $\boldsymbol{\rho_2 = 0.095238} = \alpha_2$
  - $\boldsymbol{\rho_3 = 0.095238} = \alpha_3$
  - $\boldsymbol{\rho_4 = 0.161904} = \alpha_4$

- Το μέσο πλήθος πακέτων σε κάθε υποσύστημα-ουρά M/M/1 υπολογίζεται από την επέκταση του Νόμου του Little, $N_i = \lambda_i T_i$.
- Αφού $T_i = 1/(\mu_i - \lambda_i) \Rightarrow N_i = \lambda_i / (\mu_i - \lambda_i) \Rightarrow N_i = (\lambda_i/\mu_i) / (1 - \lambda_i/\mu_i) \Rightarrow N_i = \alpha_i / (1 - \alpha_i) \Rightarrow$
- $N_i = \rho_i / (1 - \rho_i)$ (για $i = 1,\ldots,4$). Οπότε προκύπτουν οι εξής τιμές:

  $N_1 = \rho_1 / (1 - \rho_1) = \mathbf{0.135135}$

  $N_2 = \rho_2 / (1 - \rho_2) = \mathbf{0.105263}$

  $N_3 = \rho_3 / (1 - \rho_3) = \mathbf{0.105263}$

  $N_4 = \rho_4 / (1 - \rho_4) = \mathbf{0.193182}$

- Το μέσο πλήθος πακέτων σ' όλο το σύστημα των τεσσάρων ουρών, N, προκύπτει ως:
  - $N = \sum_{i=1}^{4} N_i = \mathbf{0.538843}$
- Ο χρόνος προσπέλασης του συνολικού συστήματος από εισερχόμενα πακέτα υπολογίζεται με την βοήθεια πάλι της επέκτασης του Νόμου του Little σε όλο το δίκτυο αναμονής:

$$T = N / \lambda = 0.538843 / 0.1 = \mathbf{5.38843}\ \text{s}$$

Γ. $P(0,0,0,0) = P_1(0)^* P_2(0)^* P_3(0)^* P_4(0) = (1-\alpha_1)^*(1-\alpha_2)^*(1-\alpha_3)^*(1-\alpha_4) \Rightarrow$

$P(0,0,0,0) = 0.880952^*0,904762^*0,904762^*0,838096 \Rightarrow$ **$P(0,0,0,0) = 0.604386$.**