

Μοντέλα Αναμονής

Ορισμένα απλοποιημένα μοντέλα δικτύων μπορούν να αναλυθούν με μαθηματικές μεθόδους. Τα συμπεράσματα που εξάγονται από τα αναλυτικά αποτελέσματα μπορεί είναι πολύτιμα, ακόμη και αν οι μέθοδοι δεν επιτρέπουν την αξιολόγηση λεπτομερών μοντέλων πραγματικών συστημάτων. Τα μοντέλα που χρησιμοποιούνται συνηθέστερα είναι *συστήματα ουρών αναμονής (queueing systems)*. Στην ενότητα αυτή γίνεται μία ανασκόπηση των πιο ευρέως χρησιμοποιούμενων αποτελεσμάτων.

Μία Ουρά Αναμονής FIFO

Θεωρούμε έναν καταχωρητή (buffer) εξοπλισμένο με έναν πομπό. Ο πομπός λειτουργεί με ρυθμό R bps. Εάν τα πακέτα φθάνουν στον καταχωρητή σε αρκετά μεγάλα διαστήματα μεταξύ τους, τότε κάθε φορά που ένα νέο πακέτο φθάνει στον καταχωρητή, ο καταχωρητής είναι άδειος. Στην περίπτωση αυτή, η καθυστέρηση που υφίσταται ένα πακέτο στον καταχωρητή είναι ίση με το χρόνο μετάδοσής του. Ο χρόνος μετάδοσης ενός πακέτου των L bits είναι ίσος με L/R δευτερόλεπτα.

Για παράδειγμα, εάν $R = 1$ Mbps και κάθε πακέτο αποτελείται από ακριβώς 1000 bits, τότε ο χρόνος μετάδοσης ενός πακέτου είναι ίσος με 1 ms. Εάν υποθέσουμε ότι οι χρόνοι άφιξης των πακέτων απέχουν πάντοτε τουλάχιστον 1 ms μεταξύ τους, τότε η καθυστέρηση κάθε πακέτου στον καταχωρητή είναι ίση με 1 ms. Αυτό σημαίνει ότι εάν ένα πακέτο εισέλθει στον καταχωρητή τη χρονική στιγμή T , το τελευταίο του bit αποχωρεί από τον καταχωρητή τη στιγμή $T + 1$ ms. Στο εξής θα θεωρούμε ότι ένα πακέτο φθάνει στιγμιαία στον καταχωρητή, παραδοχή που αποτελεί συνηθισμένη απλοποίηση.

Θεωρούμε και πάλι το παράδειγμά μας αλλά αυτή τη φορά υποθέτουμε ότι ο χρόνος μεταξύ διαδοχικών αφίξεων πακέτων μπορεί να είναι μικρότερος από 1 ms. Για παράδειγμα, υποθέτουμε ότι το πρώτο πακέτο φθάνει τη χρονική στιγμή 0 και το δεύτερο τη χρονική στιγμή 0,4 ms. Έστω ότι ο πομπός μεταδίδει τα πακέτα σύμφωνα με τη σειρά άφιξης. Στην περίπτωση αυτή, το τελευταίο bit του πρώτου πακέτου αποχωρεί από τον καταχωρητή τη χρονική στιγμή 1 ms. Τη στιγμή εκείνη, το δεύτερο πακέτο βρίσκεται ήδη στον καταχωρητή και αρχίζει να μεταδίδεται. Το τελευταίο bit του δεύτερου πακέτου αποχωρεί από τον καταχωρητή τη στιγμή 2 ms. Αφού το δεύτερο πακέτο έφθασε τη στιγμή 0,4 ms, ο χρόνος παραμονής του στον καταχωρητή είναι 1,6 ms. Κατά τη διάρκεια των πρώτων 0,6 ms, από τη στιγμή της άφιξής του (0,4 ms) μέχρι την αρχή της μετάδοσής του, το δεύτερο πακέτο περίμενε τη σειρά του για να μεταδοθεί. Η καθυστέρηση αυτή καλείται *καθυστέρηση αναμονής στην ουρά (queueing delay)*. Η συνολική καθυστέρηση ενός πακέτου σε έναν καταχωρητή είναι ίση με το άθροισμα της καθυστέρησης αναμονής στην ουρά και του χρόνου μετάδοσής του.

Ενώ ο χρόνος μετάδοσης ενός πακέτου εξαρτάται μόνο από το μήκος του και από το ρυθμό μετάδοσης, η καθυστέρηση αναμονής του εξαρτάται από τους χρόνους άφιξης των πακέτων στον καταχωρητή. Δηλαδή, η καθυστέρηση αναμονής στην

ουρά εξαρτάται από την κίνηση ή το φορτίο που αντιμετωπίζει ο καταχωρητής. Οι καθυστερήσεις αναμονής αυξάνουν με το φορτίο επειδή προκαλείται συμφόρηση στον καταχωρητή: ο καταχωρητής γεμίζει όταν ο πομπός δεν μπορεί να συμβαδίσει με τα πακέτα που καταφθάνουν.

Στις περισσότερες δικτυακές ζεύξεις, οι χρόνοι άφιξης των πακέτων είναι μάλλον απρόβλεπτοι. Ωστόσο, με τη μέτρηση των χρόνων άφιξης σε πολλές ζεύξεις μπορούν να κατασκευαστούν μαθηματικά μοντέλα που είναι αντιπροσωπευτικά τυπικών καταστάσεων. Τα μαθηματικά αυτά μοντέλα λαμβάνουν υπόψη τους τις ακανόνιστες διακυμάνσεις των χρόνων μεταξύ διαδοχικών αφίξεων (interarrival times). Παρόμοια μοντέλα χρησιμοποιούνται και για τα μήκη των πακέτων. Δεδομένων των μοντέλων για τους χρόνους άφιξης και τα μήκη των πακέτων, τα στατιστικά χαρακτηριστικά των καθυστερήσεων αναμονής στην ουρά και των συνολικών καθυστερήσεων στους καταχωρητές μπορούν να προβλεφθούν με τεχνικές από τη θεωρία ουρών αναμονής.

Ουρά Αναμονής M/M/1

Θεωρούμε έναν καταχωρητή εξοπλισμένο με έναν πομπό. Υποθέτουμε ότι πακέτα φθάνουν στον καταχωρητή τις χρονικές στιγμές $0 < T_1 < T_2 < T_3 < \dots$ και ότι οι χρόνοι μετάδοσης των διαδοχικών πακέτων, τα οποία μεταδίδονται με βάση την αρχή ότι αυτό που φθάνει πρώτο εξυπηρετείται πρώτο (first come-first served), είναι S_1, S_2, S_3, \dots .

Το σύστημα αυτό καλείται ουρά αναμονής M/M/1 εάν:

- Οι χρόνοι μεταξύ αφίξεων (interarrival times) $T_1, T_2 - T_1, T_3 - T_2, \dots$ είναι ανεξάρτητες τυχαίες μεταβλητές που έχουν την ίδια κατανομή με

$$P\{T_{n+1} - T_n \geq t\} = \exp\{-\lambda t\}, \quad t \geq 0. \quad (1)$$

Δηλαδή, οι χρόνοι μεταξύ αφίξεων είναι εκθετικά κατανομημένοι με ρυθμό $\lambda > 0$. Στην περίπτωση αυτή λέμε ότι οι χρόνοι άφιξης $\{T_n\}$ αποτελούν μία διαδικασία Poisson με ρυθμό λ .

- Οι χρόνοι μετάδοσης $\{S_n\}$ είναι ανεξάρτητοι και εκθετικά κατανομημένοι με ρυθμό μ . Ισοδύναμα, τα μήκη των πακέτων είναι ανεξάρτητα και εκθετικά κατανομημένα με μέση τιμή $R\mu^{-1}$ (σε bits), όπου R είναι ο ρυθμός μετάδοσης (σε bits ανά δευτερόλεπτο). Δηλαδή,

$$P\{S_n \geq t\} = \exp\{-\mu t\}, \quad t \geq 0. \quad (2)$$

Στο συμβολισμό M/M/1, το πρώτο M υποδεικνύει ότι η διαδικασία αφίξεων είναι χωρίς μνήμη (memoryless) και το δεύτερο M ότι οι χρόνοι εξυπηρέτησης είναι επίσης χωρίς μνήμη. Συμβαίνει η μόνη κατανομή χωρίς μνήμη να είναι η εκθετική. Το 1 στο M/M/1 υποδεικνύει ότι η ουρά αναμονής έχει ένα μόνο εξυπηρετητή (server).

Με τις υποθέσεις για την ουρά αναμονής M/M/1, εάν συμβολίσουμε με x_i τον αριθμό των πακέτων που είτε βρίσκονται στον καταχωρητή είτε μεταδίδονται τη

χρονική στιγμή $t > 0$, αποδεικνύεται στο Παράρτημα Β ότι, σε *στατιστική ισορροπία (statistical equilibrium)*,

$$P\{x_t = n\} = \rho^n (1 - \rho), \quad n \geq 0, \quad \text{εάν } \rho := \frac{\lambda}{\mu} < 1. \quad (3)$$

Με τον όρο στατιστική ισορροπία εννοούμε ότι το σύστημα έχει φθάσει σε μία σταθερή κατάσταση με την έννοια ότι η πιθανότητα το μήκος της ουράς αναμονής να λάβει δεδομένες τιμές δεν μεταβάλλεται με το χρόνο. Αυτό φυσικά δεν σημαίνει ότι το μήκος της ουράς παύει να εξελίσσεται. Αυτή η έννοια της ισορροπίας είναι παρόμοια με αυτό που συμβαίνει κατά την εισαγωγή ενός αερίου σε ένα άδειο μπουκάλι. Μετά από λίγο, η κατανομή των μορίων του αερίου σταθεροποιείται παρόλο που τα μόρια εξακολουθούν να κινούνται. Η πιθανότητα να βρεθεί ένας δεδομένος αριθμός μορίων σε κάποιο τμήμα του μπουκαλιού προσεγγίζει μία σταθερή τιμή.

Εάν $\lambda \geq \mu$, τότε ο μέσος αριθμός των πακέτων στον καταχωρητή είναι άπειρος. Δηλαδή, η ουρά αναμονής μεγαλώνει με το χρόνο χωρίς όριο.

Εάν $\lambda < \mu$, τότε από τη σχέση (3) συμπεραίνουμε ότι το μέσο *μήκος της ουράς αναμονής* δίνεται από τη σχέση

$$L := E\{x_t\} = \sum_{n=0}^{\infty} n P\{x_t = n\} = \frac{\rho}{1 - \rho} = \frac{\lambda}{\mu - \lambda}. \quad (4)$$

Ως εφαρμογή της σχέσης (4), θεωρήστε την παρακάτω κατάσταση. Πακέτα φθάνουν σε έναν καταχωρητή με ρυθμό μετάδοσης $R = 1$ Mbps. Τα πακέτα έχουν τυχαία μήκη L_n , τα οποία είναι ανεξάρτητα και εκθετικά κατανομημένα με μέσο μήκος ίσο με $\alpha^{-1} := 1000$ bits. Οι χρόνοι άφιξης των πακέτων αποτελούν μία διαδικασία Poisson με ρυθμό $\lambda = 800$ πακέτα/s. Θέλουμε να υπολογίσουμε το μέσο αριθμό πακέτων στον καταχωρητή, συμπεριλαμβανομένου και του πακέτου που πιθανώς να βρίσκεται υπό μετάδοση. Η ουσιαστική παρατήρηση για την εφαρμογή της σχέσης (4) είναι ότι οι χρόνοι μετάδοσης των πακέτων δίνονται από τη σχέση $S_n = L_n / R$ και επομένως είναι και αυτοί ανεξάρτητοι και εκθετικά κατανομημένοι με μέση τιμή $\mu^{-1} = 1$ ms/πακέτο. Πράγματι,

$$P\{S_n \geq t\} = P\left\{\frac{L_n}{R} \geq t\right\} = P\{L_n \geq Rt\} = \exp\{-\alpha Rt\} = \exp\{-\mu t\}$$

όπου $\mu = \alpha R = 10^{-3} \times 10^6 = 10^3$. Η τρίτη ισότητα στην παραπάνω σχέση κάνει χρήση της υπόθεσης ότι η μεταβλητή L_n είναι εκθετικά κατανομημένη με ρυθμό α . Χρησιμοποιώντας τη σχέση (4) συμπεραίνουμε ότι ο μέσος αριθμός πακέτων στον καταχωρητή είναι ίσος με $L = \lambda / (\mu - \lambda) = 800 / (1000 - 800) = 4$ πακέτα.

Στη συνέχεια χρησιμοποιούμε τη σχέση (4) για να υπολογίσουμε τη *μέση καθυστέρηση (average delay) T* ανά πακέτο στο σύστημα. Η μέση καθυστέρηση T σχετίζεται με το μέσο μήκος της ουράς με το *αποτέλεσμα του Little (Little's result)*

$$L = \lambda T. \quad (5)$$

Χρησιμοποιώντας το αποτέλεσμα αυτό συμπεραίνουμε ότι

$$T = \frac{1}{\mu - \lambda} \quad (6)$$

για την ουρά αναμονής M/M/1.

Το αποτέλεσμα του Little (5) ισχύει για μία μεγάλη κατηγορία συστημάτων ουρών αναμονής. Μία απλή ερμηνεία του αποτελέσματος αυτού μπορεί να δοθεί για συστήματα ουρών αναμονής στα οποία η σειρά εξυπηρέτησης των πελατών ακολουθεί την σειρά άφιξής τους (first come-first served). Εάν ένας τυπικός πελάτης δαπανά T μονάδες χρόνου, κατά μέσο όρο, στο σύστημα, τότε ο αριθμός των πελατών που ο τυπικός αυτός πελάτης αφήνει πίσω του κατά την αναχώρησή του είναι ίσος, με λT , κατά μέσο όρο. Πράγματι, οι πελάτες που αφήνει πίσω του είναι αυτοί που έφθασαν κατά τη διάρκεια των T μονάδων χρόνου που δαπάνησε ο τυπικός αυτός πελάτης στο σύστημα. Συνεπώς, κατά την αναχώρησή του από το δίκτυο, ένας τυπικός πελάτης αφήνει πίσω του λT πελάτες. Αυτός πρέπει να είναι ο μέσος αριθμός L πελατών στο σύστημα. Έτσι, $L = \lambda T$.

Εφαρμόζουμε τη σχέση (5) στο προηγούμενο αριθμητικό παράδειγμα για $\lambda = 800$ πακέτα/s και $\mu^{-1} = 1$ ms/πακέτο. Στην περίπτωση αυτή προκύπτει ότι

$$T = \frac{1}{\mu - \lambda} = \frac{1}{1000 - 800} = 5 \text{ ms/πακέτο.}$$

Η τιμή αυτή είναι συνεπής με το προηγούμενο συμπέρασμα ότι κατά την άφιξη ενός νέου πακέτου υπάρχουν 4 πακέτα, κατά μέσο όρο, στον καταχωρητή και κάθε πακέτο χρειάζεται 1 ms για να μεταδοθεί.

Το αποτέλεσμα του Little ισχύει και για συστήματα στα οποία η σειρά εξυπηρέτησης δεν είναι απαραίτητα η ίδια με την σειρά άφιξης, όπως εξηγείται με το παρακάτω επιχειρήμα. Έστω ότι κάθε πελάτης πληρώνει μία μονάδα κόστους για κάθε μονάδα χρόνου που δαπανά στο σύστημα. Τότε το μέσο ποσό που δαπανάται ανά μονάδα χρόνου από όλους τους πελάτες είναι ίσο με το μέσο αριθμό L των πελατών στο σύστημα σε οποιαδήποτε δεδομένη χρονική στιγμή. Από την άλλη πλευρά, κάθε πελάτης πληρώνει συνολικά T μονάδες κόστους, κατά μέσο όρο, αφού T είναι ο μέσος χρόνος που ένας πελάτης δαπανά στο σύστημα. Επομένως, αφού οι πελάτες φθάνουν με ρυθμό λ , πληρώνουν $\lambda \times T$ μονάδες κόστους ανά μονάδα χρόνου. Από τα παραπάνω έπεται η σχέση (5).

Εφαρμογή στη Στατιστική Πολυπλεξία

Τα παραπάνω αποτελέσματα για την ουρά αναμονής M/M/1 μπορούν να χρησιμοποιηθούν για να εκτιμηθεί το κέρδος που επιτυγχάνεται με τη στατιστική πολυπλεξία σε σύγκριση με την πολυπλεξία διαίρεσης χρόνου. Θεωρούμε μία γραμμή μετάδοσης με ρυθμό μετάδοσης R . Ένας αριθμός N από ροές κίνησης πακέτων, καθεμία με ρυθμό λ , πρέπει να μεταδοθούν μέσω της γραμμής αυτής. Υποθέτουμε ότι οι χρόνοι άφιξης των πακέτων αποτελούν ανεξάρτητες διαδικασίες

Poisson (με ρυθμό λ) και ότι τα μήκη των πακέτων είναι ανεξάρτητα και εκθετικά κατανομημένα με μέση τιμή L .

Η πολυπλεξία διαίρεσης χρόνου συνίσταται στη διαίρεση της χωρητικότητας του πομπού σε N κανάλια με ρυθμό R/N . Κάθε ροή κίνησης μπορεί τότε να μοντελοποιηθεί ως μία ουρά αναμονής M/M/1 με ρυθμό άφιξης λ και ρυθμό μετάδοσης $\mu := R(LN)^{-1}$ (σε πακέτα ανά δευτερόλεπτο). Η μέση καθυστέρηση ανά πακέτο δίνεται από τη σχέση (6):

$$T = \frac{1}{\mu - \lambda}. \quad (7)$$

Η στατιστική πολυπλεξία συνίσταται στην αποθήκευση των πακέτων που προέρχονται από τις N ροές σε ένα μόνο καταχωρητή και τη μετάδοσή τους ανά ένα. Η διαδικασία αφίξεων στο μοναδικό καταχωρητή έχει τώρα ρυθμό $N\lambda$. (Μπορεί να αποδειχθεί ότι η διαδικασία είναι Poisson.) Το σύστημα αυτό μπορεί να μοντελοποιηθεί ως μία ουρά αναμονής M/M/1 με ρυθμό άφιξης $N\lambda$ και ρυθμό μετάδοσης $RL^{-1} = N\mu$ πακέτα ανά δευτερόλεπτο. Η μέση καθυστέρηση δίνεται τώρα από τη σχέση

$$\frac{1}{N\mu - N\lambda} = \frac{T}{N}. \quad (8)$$

Επομένως, η μέση καθυστέρηση ανά πακέτο για το σύστημα στατιστικής πολυπλεξίας είναι ένα κλάσμα $1/N$ της καθυστέρησης για την πολυπλεξία διαίρεσης χρόνου.

Ως εφαρμογή του παραπάνω αποτελέσματος, φαντασθείτε ότι σχεδιάζετε ένα ασύρματο δίκτυο δεδομένων. Οι πομποί μπορούν να υλοποιήσουν ένα ασύρματο κανάλι με ρυθμό 128 kbps. Υπάρχουν 10 χρήστες οι οποίοι μεταδίδουν πακέτα που έχουν ανεξάρτητα και εκθετικά κατανομημένα μήκη με μέση τιμή 1000 bits. Κάθε χρήστης λαμβάνει πακέτα προς μετάδοση σύμφωνα με μία διαδικασία Poisson με ρυθμό 10 πακέτα ανά δευτερόλεπτο. Σύμφωνα με μία σχεδίαση, το κανάλι των 128 kbps διαιρείται εξίσου μεταξύ των 10 χρηστών, ώστε κάθε χρήστης να έχει στην αποκλειστική διάθεσή του ένα κανάλι των 12,8 kbps. Σύμφωνα με μία άλλη σχεδίαση, όλοι οι χρήστες μοιράζονται το κανάλι των 128 kbps. Χρησιμοποιώντας τη σχέση (7) προκύπτει ότι, στην περίπτωση της πρώτης σχεδίασης, η μέση καθυστέρηση T ανά πακέτο είναι $T = 1/(\mu - \lambda)$ με $\mu = (12,8 \text{ kbps})/(1000 \text{ bits/πακέτο}) = 12,8 \text{ πακέτα/s}$ και $\lambda = 10 \text{ πακέτα/s}$. Δηλαδή,

$$T = \frac{1}{12,8 - 10} \approx 360 \text{ ms.}$$

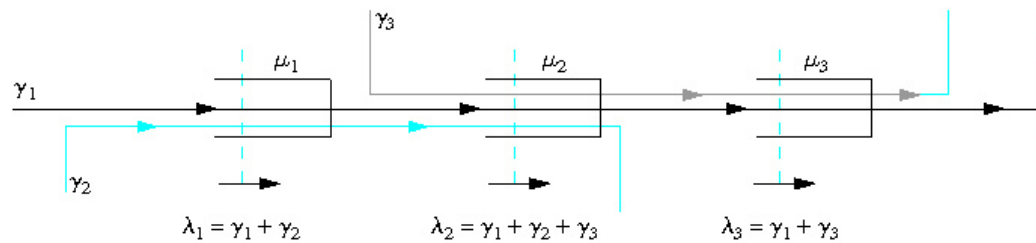
Χρησιμοποιώντας τη σχέση (8) προκύπτει ότι, στην περίπτωση της δεύτερης σχεδίασης, η μέση καθυστέρηση ανά πακέτο είναι ίση με $T/10 \approx 36 \text{ ms}$.

Δίκτυα Ουρών Αναμονής M/M/1

Τα περισσότερα δίκτυα έχουν περισσότερους από έναν καταχωρητές. Στην ενότητα αυτή θα εξηγήσουμε μερικούς απλούς τύπους οι οποίοι χρησιμοποιούνται για την εκτίμηση της μέσης καθυστέρησης ανά πακέτο σε ένα δίκτυο με πολλαπλούς καταχωρητές. Θα αρχίσουμε εξετάζοντας το δίκτυο που φαίνεται στο Σχήμα 1. Από το δίκτυο αυτό διέρχονται τρεις ροές πακέτων. Υποθέτουμε ότι οι ροές των αφίξεων είναι διαδικασίες Poisson με ρυθμούς αυτούς που αναγράφονται στο σχήμα. Υποθέτουμε επίσης ότι οι χρόνοι εξυπηρέτησης στους τρεις καταχωρητές είναι ανεξάρτητοι και εκθετικά κατανομημένοι με ρυθμούς αυτούς που υποδεικνύονται στο σχήμα.

Σχήμα 1

Ένα δίκτυο με τρεις καταχωρητές.



Είναι αξιοσημείωτο ότι το δίκτυο αυτό μπορεί να αναλυθεί αναλύοντας κάθε ουρά αναμονής ξεχωριστά. Δηλαδή, μπορεί να αποδειχθεί ότι ο μέσος αριθμός πακέτων σε κάθε καταχωρητή δίνεται από τη σχέση (5). Συγκεκριμένα, για $i = 1, 2, 3$, ο μέσος αριθμός πακέτων στον καταχωρητή i , L_i , είναι ίσος με

$$L_i = \frac{\lambda_i}{\mu_i - \lambda_i}. \quad (9)$$

Στον τύπο αυτό, λ_i είναι ο μέσος ρυθμός των πακέτων που διέρχονται από τον καταχωρητή i . Όπως φαίνεται στο Σχήμα 1, $\lambda_1 = \gamma_1 + \gamma_2$, $\lambda_2 = \gamma_1 + \gamma_2 + \gamma_3$ και $\lambda_3 = \gamma_1 + \gamma_3$.

Το εύλογο επιχείρημα που παρατέθηκε για το αποτέλεσμα του Little σε ουρές που δεν είναι FIFO (first in - first out) επεκτείνεται και στα δίκτυα ουρών αναμονής. Για την ακρίβεια, μπορεί να αποδειχθεί ότι η μέση καθυστέρηση T ανά πακέτο στο δίκτυο σχετίζεται με το μέσο αριθμό L πακέτων στο δίκτυο με τη σχέση $L = \gamma T$, όπου γ είναι ο μέσος συνολικός ρυθμός άφιξης πακέτων στο δίκτυο. Για το δίκτυο του Σχήματος 1 προκύπτει ότι

$$L = L_1 + L_2 + L_3 \text{ και } \gamma = \gamma_1 + \gamma_2 + \gamma_3. \quad (10)$$

Ως αριθμητική εφαρμογή, θεωρούμε ότι $\mu_1 = \mu_2 = \mu_3 = 1000$ πακέτα/s και ότι $\gamma_1 = \gamma_2 = \gamma_3 = 300$ πακέτα/s. Χρησιμοποιώντας τη σχέση (9) βρίσκουμε ότι

$$L_1 = \frac{\lambda_1}{\mu_1 - \lambda_1} = \frac{600}{1000 - 600} = 1,5 \text{ πακέτα.}$$

Όμοια, βρίσκουμε ότι $L_2 = 9$ πακέτα και $L_3 = 1,5$ πακέτα. Κατά συνέπεια, από τη σχέση (10) προκύπτει $L = 12$ πακέτα και, αφού $L = \gamma T$, η μέση καθυστέρηση T ανά πακέτο είναι ίση με $T = L/\gamma = 12/(900 \text{ πακέτα/s}) \approx 13,3 \text{ ms}$.

Το αποτέλεσμα που εφαρμόσαμε στο παραπάνω παράδειγμα ισχύει για πιο γενικά δίκτυα. Αποδεικνύεται ότι η μέση καθυστέρηση T ανά πακέτο σε ένα δίκτυο (σε δευτερόλεπτα) δίνεται από τον ακόλουθο τύπο:

$$T = \frac{1}{\gamma} \sum_{j=1}^J \frac{\lambda_j}{\mu_j - \lambda_j} \quad (11)$$

όπου γ είναι ο μέσος ρυθμός της ροής που εισέρχεται στο δίκτυο, λ_j είναι ο μέσος ρυθμός της ροής που διέρχεται από την ουρά j και μ_j είναι ο μέσος ρυθμός μετάδοσης της ουράς j . Όλοι αυτοί οι ρυθμοί εκφράζονται σε πακέτα ανά δευτερόλεπτο και οι ρυθμοί λ_j προκύπτουν με επίλυση των εξισώσεων που εκφράζουν τη διατήρηση των ροών διαμέσου των κόμβων.

Οι υποθέσεις που απαιτούνται ώστε να ισχύει ο τύπος (11) είναι:

1. Οι ροές αφίξεων που εισέρχονται στο δίκτυο αποτελούν ανεξάρτητες διαδικασίες Poisson.
2. Οι χρόνοι μετάδοσης των πακέτων σε όλες τις ουρές είναι ανεξάρτητοι και εκθετικά κατανομημένοι.

Στην πράξη, η υπόθεση 1 μπορεί να επιβεβαιωθεί, όχι όμως και η υπόθεση 2. Αυτό οφείλεται στο γεγονός ότι οι διαδοχικοί χρόνοι μετάδοσης ενός δεδομένου πακέτου στους διάφορους κόμβους είναι όλοι ανάλογοι προς το μήκος του πακέτου και επομένως δεν μπορούν να είναι ανεξάρτητοι. Παρόλα αυτά, πειράματα προσομοιώσεων δείχνουν ότι ο τύπος (11) παρέχει μία ικανοποιητική εκτίμηση της μέσης καθυστέρησης ανά πακέτο σε δίκτυα μεταγωγής πακέτων με αποθήκευση και προώθηση.

Ουρές Αναμονής M/G/1

Έως τώρα έχουμε θεωρήσει ότι τα μήκη των πακέτων είναι εκθετικά κατανομημένα. Μετρήσεις δείχνουν ότι το μοντέλο αυτό δεν είναι απόλυτα ακριβές. Αντίθετα, τα πακέτα τείνουν να διαχωρίζονται σε μικρά πακέτα ελέγχου και ηλεκτρονικού ταχυδρομείου και σε μεγάλα αρχεία που προκύπτουν από μεταφορές αρχείων. Εξαιτίας του γεγονότος αυτού, χρειάζεται να εξετάσουμε την επίδραση της κατανομής του μήκους των πακέτων στη συμπεριφορά των ουρών αναμονής.

Θεωρούμε και πάλι την ουρά αναμονής M/M/1 αλλά με την υπόθεση ότι οι διαδοχικοί χρόνοι μετάδοσης πακέτων S_n είναι ανεξάρτητοι και έχουν κάποια κοινή κατανομή που δεν είναι απαραίτητα εκθετική. Το σύστημα που προκύπτει καλείται ουρά αναμονής M/G/1. Στο συμβολισμό M/G/1, το M υποδεικνύει ότι η διαδικασία αφίξεων είναι χωρίς μνήμη, δηλαδή ότι έχει εκθετικά κατανομημένους χρόνους μεταξύ αφίξεων και, επομένως, είναι Poisson. Το G υποδεικνύει ότι οι χρόνοι

εξυπηρέτησης έχουν μία γενική κατανομή. Το 1 υπενθυμίζει ότι υπάρχει ένας μοναδικός εξυπηρετητής.

Αποδεικνύουμε στο Παράρτημα Β ότι η μέση καθυστέρηση ανά πακέτο που διέρχεται από το σύστημα δίνεται από τη σχέση

$$T = \frac{1}{\mu} + \frac{\lambda E\{S_n^2\}}{2(1-\rho)} = \frac{1}{\mu} + \frac{\lambda(\sigma^2 + \mu^{-2})}{2(1-\rho)} \quad (12)$$

όπου μ^{-1} ο μέσος χρόνος μετάδοσης, $\rho = \lambda/\mu$ και σ^2 η μεταβλητότητα των χρόνων μετάδοσης. Δηλαδή, $\sigma^2 = E\{(S_n - \mu^{-1})^2\}$. Η σχέση (12) δείχνει ότι η μέση καθυστέρηση αυξάνει με τη μεταβλητότητα του μήκους των πακέτων.

Για παράδειγμα, εάν οι χρόνοι μετάδοσης είναι εκθετικά κατανομημένοι με μέση τιμή μ^{-1} , τότε $\sigma^2 = \mu^{-2}$, οπότε η σχέση (12) καταλήγει στη σχέση (6). Ως άλλο παράδειγμα, εάν όλα τα πακέτα έχουν το ίδιο μήκος, τότε $\sigma^2 = 0$ και η μέση καθυστέρηση των πακέτων είναι ίση με $(1 - \rho/2)$ φορές την καθυστέρηση σε μία ουρά αναμονής M/M/1.

Ως ένα επιπλέον παράδειγμα, υποθέτουμε ότι οι χρόνοι μετάδοσης των πακέτων S_n είναι τέτοιοι ώστε

$$P(S_n = a) = p = 1 - P(S_n = b)$$

όπου $p \in (0, 1)$ και $0 < a < b$ είναι δύο δεδομένες διάρκειες. Δηλαδή, το μήκος των πακέτων έχει μόνο δύο δυνατές τιμές. Στην περίπτωση αυτή, $\mu^{-1} = pa + (1-p)b$ και $E\{S_n^2\} = pa^2 + (1-p)b^2$. Ανάλογα, βρίσκουμε ότι

$$T = \frac{1}{pa + (1-p)b} + \frac{\lambda(pa^2 + (1-p)b^2)}{2[1 - \lambda(pa + (1-p)b)]}$$

Για παράδειγμα, εάν $\lambda = 300$ πακέτα/s, $p = 0,8$, $a = 0,8$ ms και $b = 10$ ms, τότε προκύπτει ότι $T \approx 17,4$ ms.

Προσοχή

Το αποτέλεσμα της σχέσης (12) δείχνει ότι πρέπει να είμαστε προσεκτικοί όταν κάνουμε υποθέσεις για την κατανομή των μηκών των πακέτων. Το αποτέλεσμα αυτό δείχνει ότι οι μέσες τιμές των χρόνων εξυπηρέτησης και των χρόνων μεταξύ αφίξεων δεν επαρκούν για την πρόβλεψη των μέσων καθυστερήσεων. Η μεταβλητότητα των χρόνων εξυπηρέτησης έχει σημαντική επίδραση στη μέση καθυστέρηση.

Παρόμοια προσοχή πρέπει να επιδεικνύεται ως προς τη διαδικασία άφιξης των πακέτων. Για παράδειγμα, εάν υποθέσουμε ότι πακέτα φθάνουν σε μία ουρά αναμονής κάθε λ^{-1} μονάδες χρόνου ακριβώς και ότι οι χρόνοι εξυπηρέτησής τους είναι σταθεροί και ίσοι με $\mu^{-1} < \lambda^{-1}$, τότε κάθε πακέτο φθάνει αφού έχουν μεταδοθεί τα προηγούμενα πακέτα, οπότε δεν υπάρχει αναμονή στην ουρά. Κατά συνέπεια, η καθυστέρηση για κάθε πακέτο είναι ίση με το χρόνο μετάδοσης, δηλαδή, ίση με μ^{-1} . Συγκρίνοντας το αποτέλεσμα αυτό με τη σχέση (12),

επιβεβαιώνουμε ότι οι κατανομές των χρόνων μεταξύ αφίξεων και των χρόνων εξυπηρέτησης επηρεάζουν σημαντικά τις καθυστερήσεις.

Θα πρέπει λοιπόν να είμαστε προσεκτικοί όταν προβαίνουμε σε βιαστικές υποθέσεις, ακόμη και όταν αυτές μας διευκολύνουν. Μπορεί να είναι πολύ δελεαστικό να υποθέσουμε ότι κάποιες διαδικασίες είναι Poisson ώστε να είμαστε σε θέση να διεξάγουμε κάποια ανάλυση. Ωστόσο, οι μηχανικοί δικτύων δεν πρέπει να εμπιστεύονται τυφλά τα αποτελέσματα μίας ανάλυσης βασισμένης σε τέτοιες αυθαίρετες υποθέσεις. Εάν μία πιο ακριβής ανάλυση δεν είναι εφικτή και εφόσον οι υποθέσεις δεν μπορούν να επικυρωθούν, τότε είναι λιγότερο παραπλανητικό να παραδεχθούμε το γεγονός αυτό και να καταφύγουμε σε μία προσεκτική προσομοίωση από το να κάνουμε αυθαίρετες υποθέσεις χάριν ευκολίας.

Προσεκτικές μετρήσεις σε ροές κίνησης δείχνουν ότι απλά μοντέλα μπορούν να περιγράψουν τη συμπεριφορά τους για χρονικές περιόδους με ικανοποιητικά μεγάλη διάρκεια αν και είναι ανακριβή για μεγαλύτερες περιόδους παρατήρησης. Επιπρόσθετα, αναλυτικές μελέτες και προσομοιώσεις δείχνουν ότι τα μοντέλα αυτά παρέχουν ικανοποιητικές προβλέψεις για τις καθυστερήσεις και τα αποθέματα στους καταχωρητές. Ωστόσο, τα κατάλληλα μοντέλα είναι πιο πολύπλοκα από διαδικασίες Poisson και ανεξάρτητους χρόνους εξυπηρέτησης. Το συμπέρασμα των μελετών αυτών είναι ότι η ανάλυση παρέχει ισχυρά εργαλεία για την εξέταση της λειτουργίας των δικτύων. Τα εργαλεία αυτά μας επιτρέπουν να κατανοήσουμε τον τρόπο με τον οποίο τα διάφορα χαρακτηριστικά των δικτύων επηρεάζουν την απόδοσή τους. Για παράδειγμα, από το απλό μοντέλο M/G/1 κατανοούμε τη σπουδαιότητα της μεταβλητότητας του χρόνου εξυπηρέτησης. Πιο ολοκληρωμένες αναλυτικές μελέτες επαυξάνουν μία τέτοια ποιοτική κατανόηση και παρέχουν πληροφορίες χρήσιμες στην εφαρμογή των μεθοδολογιών μετρήσεων και προσομοιώσεων καθώς και στη διεξαγωγή πειραμάτων.

Ουρές Αναμονής με Διακοπές

Θεωρούμε την ακόλουθη παραλλαγή της ουράς αναμονής M/G/1. Υποθέτουμε ότι αμέσως μόλις ο πομπός αδειάσει την ουρά απενεργοποιείται για κάποιο τυχαίο χρονικό διάστημα, το οποίο καλείται *περίοδος διακοπών*. Υποθέτουμε επίσης ότι οι διαδοχικές διακοπές έχουν διάρκειες που είναι ανεξάρτητες μεταξύ τους, ανεξάρτητες της προηγούμενης εξέλιξης του συστήματος και έχουν κοινή κατανομή.

Με αυτές τις υποθέσεις μπορεί να αποδειχθεί ότι η μέση καθυστέρηση ανά πακέτο δίνεται από τη σχέση

$$T = T_0 + \frac{\nu}{2} E\{V^2\} \quad (13)$$

όπου T_0 είναι η μέση καθυστέρηση σε μία ουρά αναμονής M/G/1 χωρίς διακοπές, V συμβολίζει μία τυπική διάρκεια διακοπών και ν^{-1} είναι η μέση διάρκεια διακοπών.

Για παράδειγμα, εάν οι διακοπές έχουν σταθερή διάρκεια ίση με A , τότε από τη σχέση (13) προκύπτει ότι $T = T_0 + A/2$. Αυτή η συγκεκριμένη περίπτωση δεν προκαλεί μεγάλη έκπληξη. Πράγματι, ένα πακέτο που φθάνει στην ουρά ενδέχεται να βρεί άλλα πακέτα, των οποίων η μετάδοση έχει καθυστερήσει κατά το μισό του χρόνου διακοπών $A/2$, κατά μέσο όρο, ενώ περιμένουν να επανενεργοποιηθεί ο πομπός.

Ενδιαφέρον παρουσιάζει η παρατήρηση ότι, εάν οι περίοδοι διακοπών δεν είναι όλες σταθερές, τότε από τη σχέση (13) διαπιστώνεται ότι η επίδραση τους στη μέση καθυστέρηση είναι μεγαλύτερη από τη μέση διάρκεια διακοπών. (Αναλυτικά, $E\{V^2\} > (E\{V\})^2 = v^{-2}$, οπότε $T > T_0 + v^{-1}/2$.) Διαισθητικά, είναι πιθανότερο τα πακέτα να καθυστερήσουν εξαιτίας των διακοπών με μεγαλύτερη διάρκεια αφού αυτές επηρεάζουν περισσότερα πακέτα από ότι οι διακοπές μικρότερης διάρκειας.

Η ουρά αναμονής με διακοπές αποτελεί μία εισαγωγή στην ανάλυση ενός καταχωρητή ο οποίος μοιράζεται τον πομπό μαζί με άλλους καταχωρητές, που εξετάζεται παρακάτω στην Ενότητα “Συστήματα Κυκλικής Εξυπηρέτησης”.

Συστήματα Προτεραιότητων

Θεωρούμε συστήματα μετάδοσης με δύο είδη πακέτων: φωνής και δεδομένων. Υποθέτουμε ότι οι αφίξεις ακολουθούν δύο ανεξάρτητες Poisson διαδικασίες με ρυθμούς λ_1 για τα πακέτα φωνής και λ_2 για τα πακέτα δεδομένων. Υποθέτουμε επίσης ότι οι χρόνοι μετάδοσης των πακέτων φωνής είναι όλοι ανεξάρτητοι και κατανομημένοι όπως η τυχαία μεταβλητή S_1 και ότι οι χρόνοι μετάδοσης των πακέτων δεδομένων είναι ανεξάρτητοι και κατανομημένοι όπως η τυχαία μεταβλητή S_2 .

Στα πακέτα φωνής δίνεται *προτεραιότητα* έναντι των πακέτων δεδομένων. Δηλαδή, όποτε ολοκληρώνεται η μετάδοση ενός πακέτου, ο πομπός αρχίζει τη μετάδοση ενός πακέτου φωνής εφόσον υπάρχει ένα στον καταχωρητή, διαφορετικά, αρχίζει τη μετάδοση ενός πακέτου δεδομένων, εάν υπάρχει.

Στο Παράρτημα Β αποδεικνύεται ότι η μέση καθυστέρηση ανά πακέτο φωνής δίνεται από τη σχέση $\mu_1^{-1} + W_1$ και η μέση καθυστέρηση ανά πακέτο δεδομένων από την $\mu_2^{-1} + W_2$, όπου $\mu_i^{-1} = E\{S_i\}$ για $i = 1, 2$, και

$$W_1 = \frac{\sum_{i=1}^2 \lambda_i E\{S_i^2\}}{2(1 - \rho_1)} \quad \text{και} \quad W_2 = \frac{\sum_{i=1}^2 \lambda_i E\{S_i^2\}}{2(1 - \rho_1)(1 - \rho_1 - \rho_2)} \quad (14)$$

όπου $\rho_i := \lambda_i \mu_i^{-1}$ για $i = 1, 2$.

Για παράδειγμα, έστω ότι $\lambda_1 = 300$, $\lambda_2 = 10$, $S_1 \equiv 2$ ms και S_2 είναι εκθετικά κατανομημένη με μέση τιμή 20 ms. Βρίσκουμε ότι $E\{S_2^2\} = 8 \times 10^{-4}$ s². Με αντικατάσταση στις σχέσεις (14) προκύπτει $W_1 = 11,5$ ms και $W_2 = 57,5$ ms. Εφαρμόζοντας τις σχέσεις (14) για άλλη μία φορά, μπορούμε να επαληθεύσουμε ότι η αντιστροφή των προτεραιοτήτων έχει ως αποτέλεσμα $W_1 = 28,75$ ms και $W_2 =$

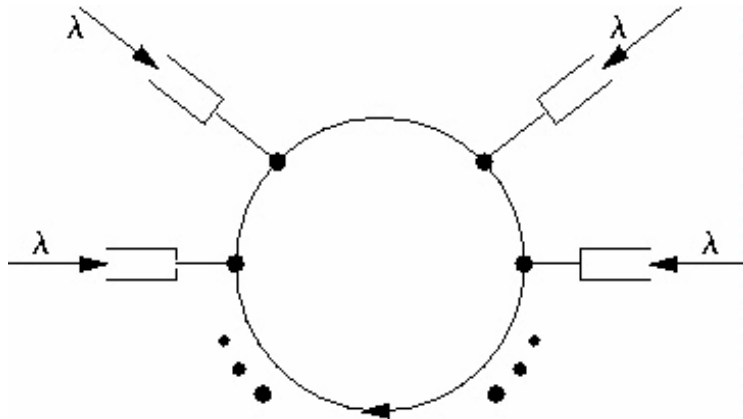
5,75 ms. Από τη σύγκριση των δύο παραπάνω αποτελεσμάτων με ανεστραμμένες προτεραιότητες φαίνεται καθαρά η επίδραση των προτεραιοτήτων στο μέσο χρόνο αναμονής στην ουρά.

Συστήματα Κυκλικής Εξυπηρέτησης

Θεωρούμε το μοντέλο του δικτύου δακτυλίου με κουπόνι που απεικονίζεται στο Σχήμα 2. Στο μοντέλο αυτό τα πακέτα φθάνουν σε κάθε σταθμό ως διαδικασίες αφίξεων Poisson με τον ίδιο ρυθμό λ . Οι χρόνοι μετάδοσης των πακέτων είναι ανεξάρτητοι και κατανέμονται όπως η τυχαία μεταβλητή S . Το κουπόνι περιφέρεται στο δακτύλιο έως ότου περιέλθει στην κατοχή ενός σταθμού του οποίου η ουρά δεν είναι άδεια. Τη στιγμή εκείνη ο σταθμός μπορεί να μεταδώσει ένα πακέτο. Μετά τη μετάδοση απελευθερώνει το κουπόνι. Αυτό είναι το πρωτόκολλο που εξετάσαμε στην Ενότητα 4.3. Ο χρόνος περιφοράς του κουπονιού στο δακτύλιο, όταν αυτό δεν περιέρχεται στην κατοχή κανενός σταθμού, είναι ίσος με R .

Σχήμα 2

Δίκτυο δακτυλίου με κουπόνι. Ο χρόνος περιφοράς του κουπονιού όταν οι ουρές είναι άδειες είναι R .



Το πρόβλημα που τίθεται είναι ο υπολογισμός της μέσης καθυστέρησης ανά πακέτο. Έχει αποδειχθεί ότι η μέση καθυστέρηση είναι ίση με $E\{S\} + W = \mu^{-1} + W$, όπου

$$W = \frac{N\lambda E\{S^2\} + R[1 + \rho N^{-1}]}{2[1 - \rho - \lambda R]} \quad (15)$$

Στην παραπάνω έκφραση, $\rho := N\lambda\mu^{-1}$. Η σχέση αυτή ισχύει με την προϋπόθεση ότι $\lambda R < 1 - \rho$, διαφορετικά το σύστημα είναι ασταθές.

Ας σημειωθεί ότι εάν ο χρόνος R (ο χρόνος περιφοράς του κουπονιού) είναι αμελητέος τότε από τη σχέση αυτή προκύπτει

$$W = \frac{N\lambda E\{S^2\}}{2(1 - \rho)},$$

το οποίο είναι ίδιο με το χρόνο αναμονής σε μία ουρά M/G/1. Το γεγονός αυτό δεν πρέπει να μας εκπλήσσει αφού το σύστημα συμπεριφέρεται ακριβώς όπως μία ουρά αναμονής M/G/1 με τη διαφορά ότι η σειρά εξυπηρέτησης δεν είναι απαραίτητα ίδια με τη σειρά άφιξης. Η τροποποίηση αυτή δεν μπορεί να μεταβάλλει το μέσο αριθμό πακέτων στο σύστημα και, επομένως, σύμφωνα με το αποτέλεσμα του Little, δεν μπορεί να επηρεάσει τη μέση καθυστέρηση.

Η συνθήκη ευστάθειας $\lambda R < 1 - \rho$ μπορεί να γραφεί ως $E\{NS + R\} < \lambda^{-1}$. Η συνθήκη αυτή μπορεί να ερμηνευτεί εάν παρατηρήσουμε ότι όταν όλες οι ουρές δεν είναι άδειες, ο μέσος χρόνος μεταξύ δύο διαδοχικών επισκέψεων του κουπονιού σε ένα σταθμό είναι $E\{NS + R\}$. Αυτός ο μέσος χρόνος θα πρέπει να είναι μικρότερος από το μέσο χρόνο μεταξύ αφίξεων λ^{-1} σε ένα σταθμό.