

# ΣΥΣΤΗΜΑΤΑ ΑΝΑΜΟΝΗΣ

## Queuing Systems

**Εφαρμογές Κλειστών Δικτύων Ουρών Markov:**

- 1. Ανάλυση Window Flow Control σε Δίκτυα Υπολογιστών**
- 2. Αξιολόγηση Συστημάτων Πολύ-προγραμματισμού (Multitasking)**
- 3. Ανάλυση Μέσης Τιμής (Mean Value Analysis – MVA)**
- 4. Προσομοίωση Κλειστού Δικτύου Markov**

**Γενίκευση Μοντέλων Μορφής Γινομένου (BCMP)**

**Βασίλης Μάγκλαρης**

[maglaris@netmode.ntua.gr](mailto:maglaris@netmode.ntua.gr)

**22/5/2019**

# ΚΛΕΙΣΤΟ ΔΙΚΤΥΟ ΔΥΟ ΕΚΘΕΤΙΚΩΝ ΟΥΡΩΝ (επανάληψη)

## Θεώρημα Gordon-Newell για $M = 2$ Ουρές, $N = 3$ Πελάτες

$$X_1 \mu_1 = X_2 \mu_2$$

$$X_1 = 1, X_2 = \mu_1 / \mu_2 = \alpha$$

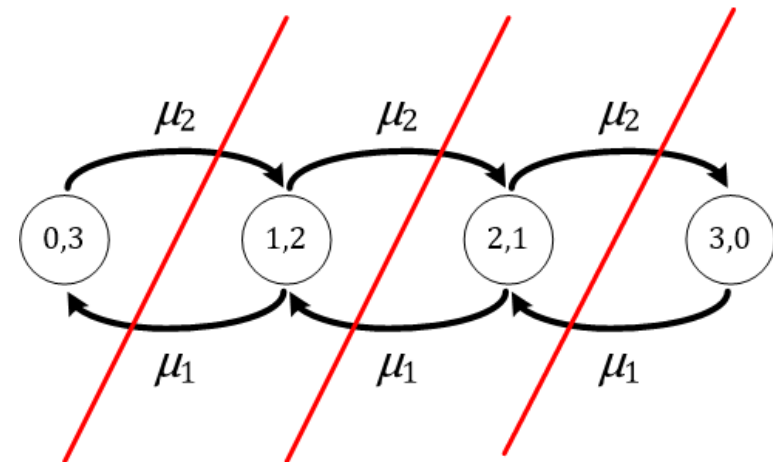
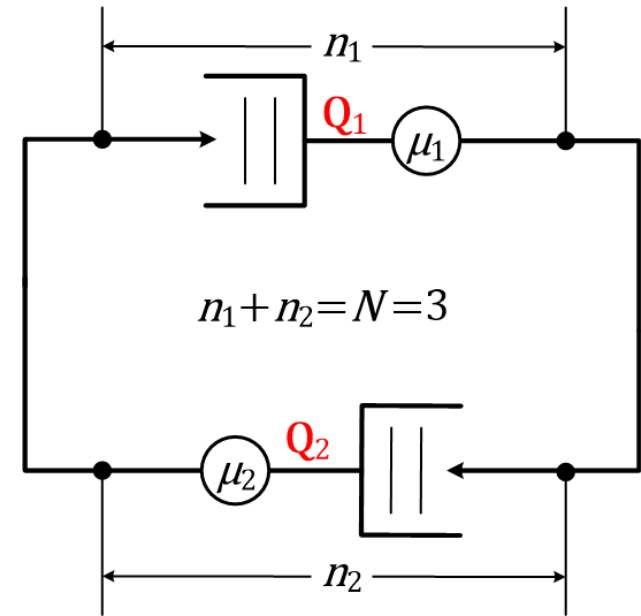
- $P(0,3) = X_2^3 / G(3) = \alpha^3 / G(3)$
- $P(1,2) = X_2^2 / G(3) = \alpha^2 / G(3)$
- $P(2,1) = X_2 / G(3) = \alpha / G(3)$
- $P(3,0) = 1 / G(3)$
- $1 / G(3) + \alpha / G(3) + \alpha^2 / G(3) + \alpha^3 / G(3) = 1$

$$\text{Άρα: } G(3) = 1 / (1 + \alpha + \alpha^2 + \alpha^3)$$

$$\gamma = \mu_2 [1 - P(3,0)] = \mu_2 [1 - 1 / G(3)]$$

$$E(T_j) = E(n_j) / \gamma$$

- $E(n_1) = P(1,2) + 2P(2,1) + 3P(3,0) = \frac{[a^2 + 2a + 3]}{G(3)}$
- $E(n_2) = P(2,1) + 2P(1,2) + 3P(0,3) = \frac{[a + 2a^2 + 3a^3]}{G(3)}$
- $E(n_1) + E(n_2) = N = 3$



# ΚΛΕΙΣΤΟ ΔΙΚΤΥΟ ΔΥΟ ΕΚΘΕΤΙΚΩΝ ΟΥΡΩΝ (επανάληψη)

## Αλγόριθμος Buzen για $M = 2$ Ουρές, $N = 3$ Πελάτες

Με βάση τον Επαναληπτικό Αλγόριθμο του **Buzen**

και  $X_1 = 1, X_2 = \mu_1/\mu_2 = a$  ισχύει ότι:

$$g(n, m) = g(n, m - 1) + X_m g(n - 1, m)$$

$$G(N) = g(N, M)$$

Προκύπτει :

$$G(1) = 1 + a$$

$$G(2) = 1 + a + a^2$$

$$G(3) = 1 + a + a^2 + a^3$$

Πίνακας Τιμών  $g(n, m)$

$n$	$X_1$	$X_2$
0	1	1
1	1	$1 + a$
2	1	$1 + a + a^2$
3	1	$1 + a + a^2 + a^3$

Η χρησιμοποίηση της ουράς  $Q_1$  είναι

$$\begin{aligned} P(n_i \geq 1) &= P(1,2) + P(2,1) + P(3,0) = 1 - P(0,3) = X_i G(N - 1)/G(N) = X_1 G(2)/G(3) = \\ &= \frac{1+a+a^2}{1+a+a^2+a^3} \quad (\text{όπως και στη 2}^\eta \text{ διαφάνεια}) \end{aligned}$$

Επίσης:

$$P(0,3) = a^3/G(3)$$

$$P(1,2) = a^2/G(3)$$

$$P(2,1) = a/G(3)$$

$$P(3,0) = 1/G(3)$$

$$E[n_i] = \sum_{k=1}^N X_i^k \frac{G(N - k)}{G(N)}$$

$$E(n_1) = \sum_{k=1}^3 \frac{G(3 - k)}{G(3)} = \frac{a^2 + 2a + 3}{1 + a + a^2 + a^3}$$

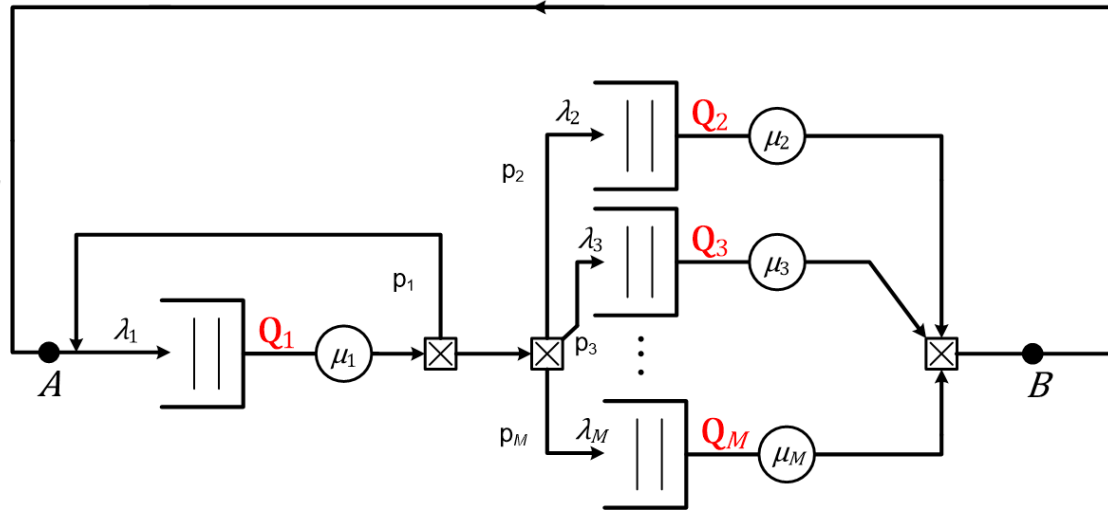
$$E(n_2) = \sum_{k=1}^3 X_2^k \frac{G(N - k)}{G(N)} = 3 - E(n_1) = \frac{a + 2a^2 + 3a^3}{1 + a + a^2 + a^3}$$

# ΚΛΕΙΣΤΟ ΔΙΚΤΥΟ ΕΚΘΕΤΙΚΩΝ ΟΥΡΩΝ (1/2) (επανάληψη)

Βασισμένο στο Παράδειγμα του *Jeffrey Buzen*, "Computational Algorithms for Closed Queuing Networks with Exponential Servers", Communications of the ACM 16 (9), Sept. 1973

Κλειστό δίκτυο  $M$  εκθετικών ουρών  $Q_1$  (CPU),  $Q_2, \dots, Q_M$  (I/O)

Παράλληλη Επεξεργασία  $N$  προγραμμάτων (εντολών) με ανακύκλωση στη CPU (πιθανότητα  $p_{11} = p_1$ ), επιλογή Υποσυστήματος I/O (με πιθανότητες  $p_{12} = p_2, p_{13} = p_3, \dots, p_{1M} = p_M$ ) και απάντηση – δημιουργία νέας εντολής (εξωτερική ανάδραση).



Εφαρμογή Αλγορίθμου *Buzen* για  $N = 1, 2, 3, 4$  πελάτες (παράλληλα προγράμματα) και  $M = 3$  ουρές  $Q_1, Q_2, Q_3$

$$\mu_1 X_1 = p_1 \mu_1 X_1 + \mu_2 X_2 + \mu_3 X_3$$

$$\mu_2 X_2 = p_2 \mu_1 X_1$$

$$\mu_3 X_3 = p_3 \mu_1 X_1$$

$$\text{Με } \mu_1 = \frac{1}{28} \text{ msec}^{-1}, \mu_2 = \frac{1}{40} \text{ msec}^{-1}, \mu_3 = \frac{1}{280} \text{ msec}^{-1},$$

$p_1 = 0.1, p_2 = 0.7, p_3 = 0.2$  και  $X_1 = 1$  προκύπτει πως  $X_2 = 1, X_3 = 2$

Ο αναδρομικός τύπος  $g(n, m) = g(n, m - 1) + X_m g(n - 1, m)$  δίνει τον πίνακα δεξιά

Οι σταθερές  $G(N) = g(N, 3)$  αντιστοιχούν σε  $N = 1, 2, 3, 4$  προγράμματα  
 $G(1) = 4, \quad G(2) = 11, \quad G(3) = 26, \quad G(4) = 57$

Πίνακας Τιμών  $g(n, m)$

$n$	$X_1$	$X_2$	$X_3$
0	1	1	1
1	1	2	4
2	1	3	11
3	1	4	26
4	1	5	57

# ΚΛΕΙΣΤΟ ΔΙΚΤΥΟ ΕΚΘΕΤΙΚΩΝ ΟΥΡΩΝ (2/2) (επανάληψη)

Βασισμένο στο Παράδειγμα του *Jeffrey Buzen*, "Computational Algorithms for Closed Queuing Networks with Exponential Servers", *Communications of the ACM* 16 (9), Sept. 1973

Οι αντίστοιχοι βαθμοί  
χρησιμοποίησης της CPU ( $Q_1$ )  
 $U_1 = G(N - 1)/G(N)$  είναι:

$N$	1	2	3	4
$U_1$	1/4	4/11	11/26	26/57

Η ρυθμαπόδοση του συστήματος είναι

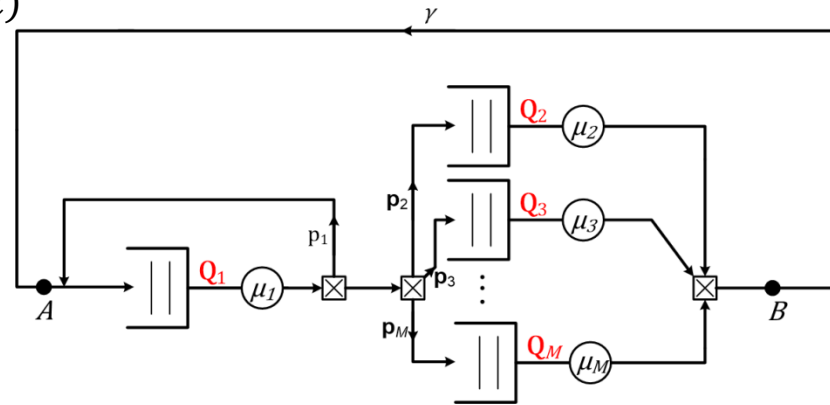
$$\begin{aligned} \gamma &= \mu_2 p(n_2 \geq 1) + \mu_3 p(n_3 \geq 1) \\ &= \mu_2 X_2 G(N - 1)/G(N) + \mu_3 X_3 G(N - 1)/G(N) \\ &= (\mu_2 X_2 + \mu_3 X_3) G(N - 1)/G(N) \end{aligned}$$

Οι αντίστοιχες τιμές σε προγράμματα/sec είναι:

$N$	1	2	3	4
$\gamma$	8.04	11.69	13	14.66

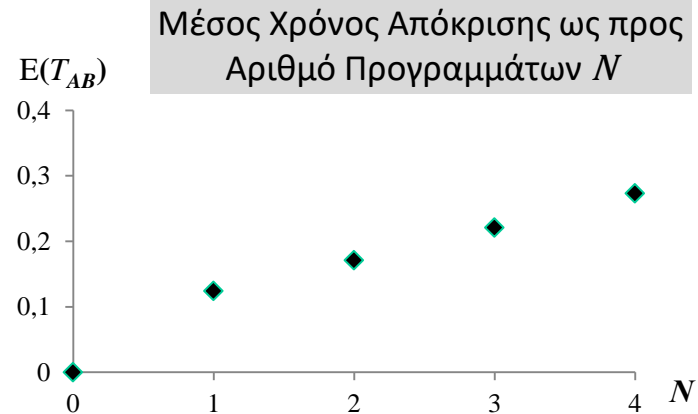
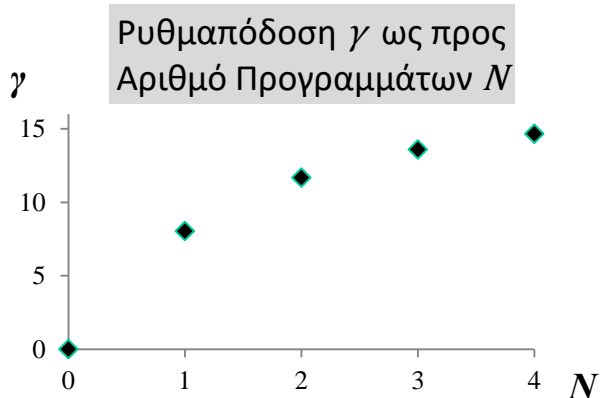
Πίνακας Τιμών  $g(n, m)$

$n$	$X_1$	$X_2$	$X_3$
0	1	1	1
1	1	2	4
2	1	3	11
3	1	4	26
4	1	5	57



Ο μέσος χρόνος απόκρισης είναι  $E(T_{AB}) = N/\gamma$   
Οι αντίστοιχες τιμές σε sec είναι:

$N$	1	2	3	4
$E(T_{AB})$	0.124	0.171	0.221	0.273



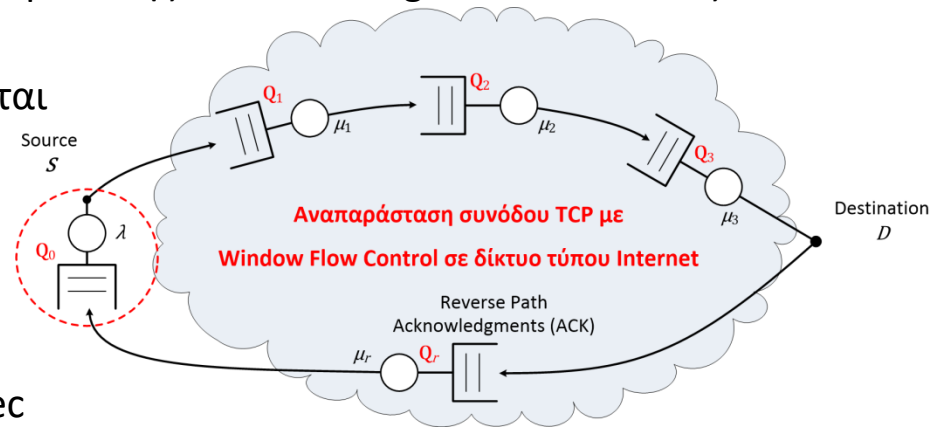
# ΕΛΕΓΧΟΣ ΡΟΗΣ ΑΠΟ ΑΚΡΟ ΣΕ ΑΚΡΟ ΣΤΟ INTERNET (1/10)

## End-to-End Window Flow Control TCP, Μέγεθος Παραθύρου $W$ (επανάληψη)

Αναπαράσταση Μηχανισμού Ελέγχου Ροής Παραθύρου (**Window Flow Control**) μέσω Κλειστού Δικτύου  $M$  ανεξαρτήτων εκθετικών ουρών και  $W$  πελατών. Η μορφή των πελατών εναλλάσσεται ανάμεσα σε πακέτα δεδομένων, μηνύματα επιβεβαίωσης – acknowledgments και άδειες εκπομπής – tokens

Το κλειστό δίκτυο του παραδείγματος αποτελείται από  $M = 5$  υποσυστήματα:

- $Q_0$ : Αποθηκεύει τα Tokens στην πηγή (Source) με τον μηχανισμό Window Flow Control και αποστέλλει στον προορισμό (Destination) νέα πακέτα ανά χρονικά διαστήματα μέσης τιμής  $1/\lambda$  sec (μοντέλο δημιουργίας κίνησης  $\lambda$  πακέτα/sec) εφόσον υπάρχουν διαθέσιμα tokens στην  $Q_0$
- $Q_1, Q_2, Q_3$ : Ενδιάμεσοι δικτυακοί κόμβοι μεταγωγής πακέτου με μέσους εκθετικούς ρυθμούς  $\mu_1, \mu_2, \mu_3$  πακέτα/sec
- $Q_r$ : Ισοδύναμο μοντέλο καθυστέρησης για την δημιουργία και μεταβίβαση μηνυμάτων επιβεβαίωσης – ACK σαν ανεξάρτητη ουρά με μέσο εκθετικό ρυθμό  $\mu_r$  πακέτα/sec (θεωρούμε κατά προσέγγιση  $1/\mu_r \cong 1/\mu_1 + 1/\mu_2 + 1/\mu_3$  sec)

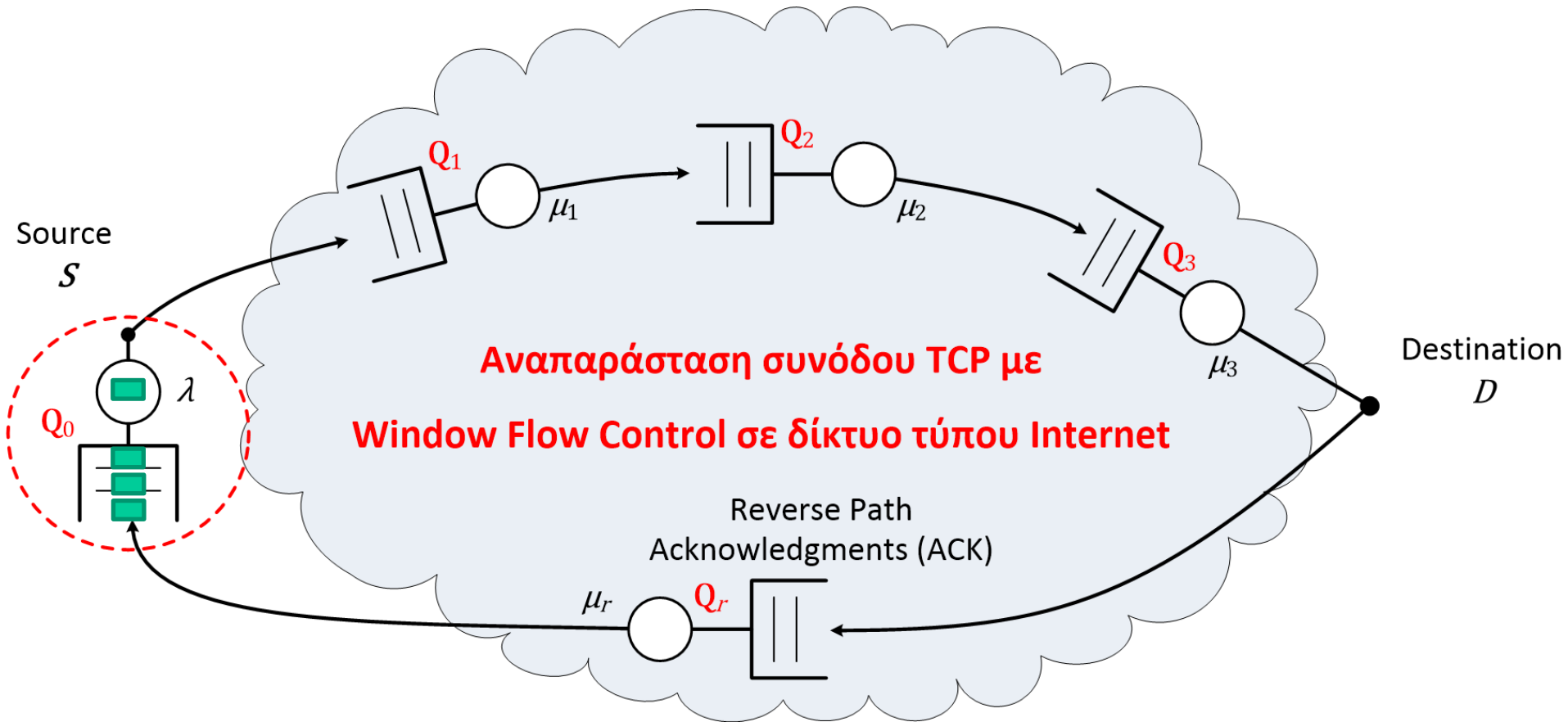


Στο κλειστό δίκτυο υπάρχουν ανά πάσα στιγμή  $W \leq 8$  πελάτες που αντιστοιχούν στο μέγεθος παραθύρου – Window Size:  $W = n_0 + n_1 + n_2 + n_3 + n_r$

Θεωρούμε πως ισχύουν οι παραδοχές για μορφή γινομένου του θεωρήματος **Gordon – Newell** και εφαρμόζουμε τον Αλγόριθμο του **Buzen** για  $W = 1, \dots, 8$  πελάτες που κυκλοφορούν σε  $M = 5$  ουρές, ως προς την ρυθμαπόδοση  $\gamma$  και την μέση καθυστέρηση πακέτου από άκρο σε άκρο ( $S$  σε  $D$ ) στο δίκτυο  $E(T_{S,D}) = [E(n_1) + E(n_2) + E(n_3)]/\gamma$

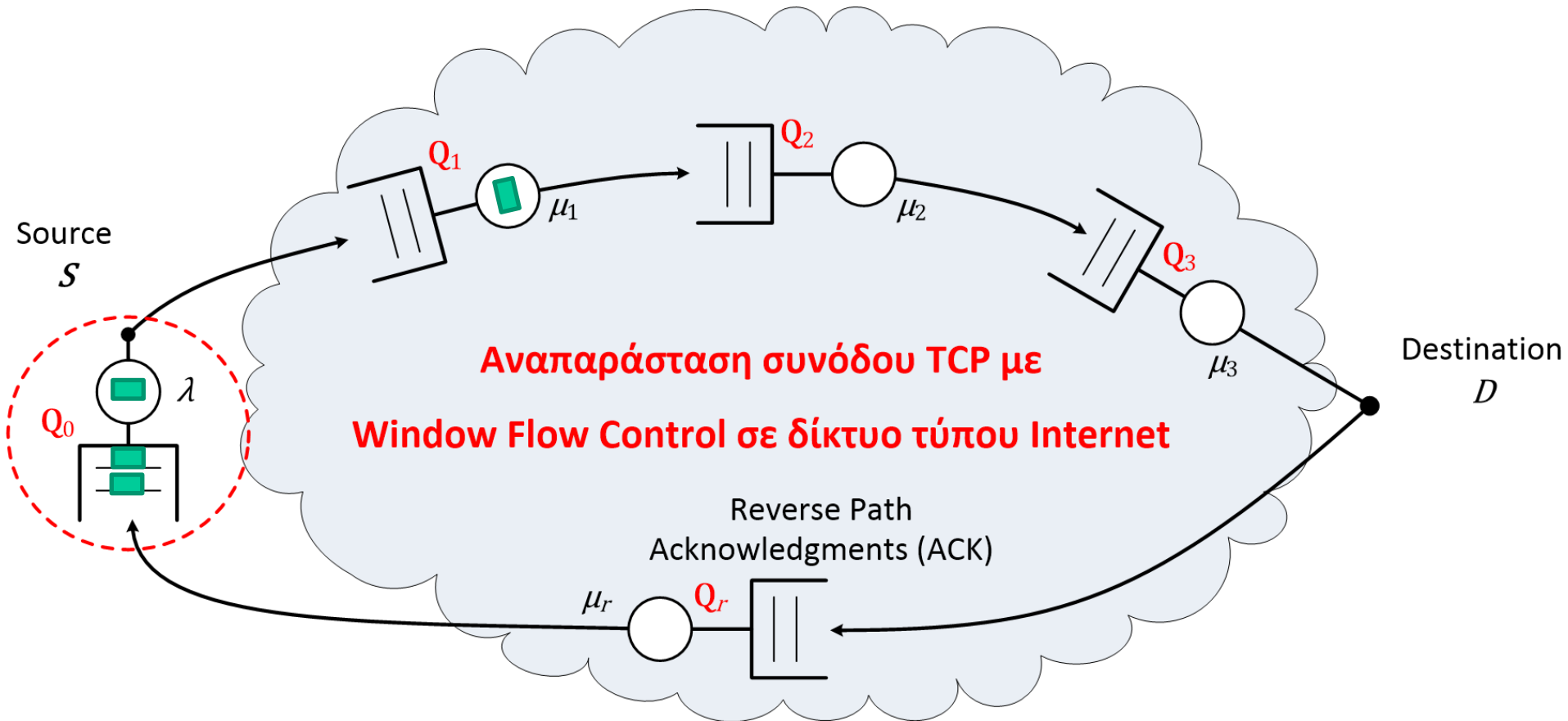
# ΕΛΕΓΧΟΣ ΡΟΗΣ ΑΠΟ ΑΚΡΟ ΣΕ ΑΚΡΟ ΣΤΟ INTERNET (2/10)

End-to-End Window Flow Control TCP, Μέγεθος Παραθύρου  $W$  (επανάληψη)



# ΕΛΕΓΧΟΣ ΡΟΗΣ ΑΠΟ ΑΚΡΟ ΣΕ ΑΚΡΟ ΣΤΟ INTERNET (3/10)

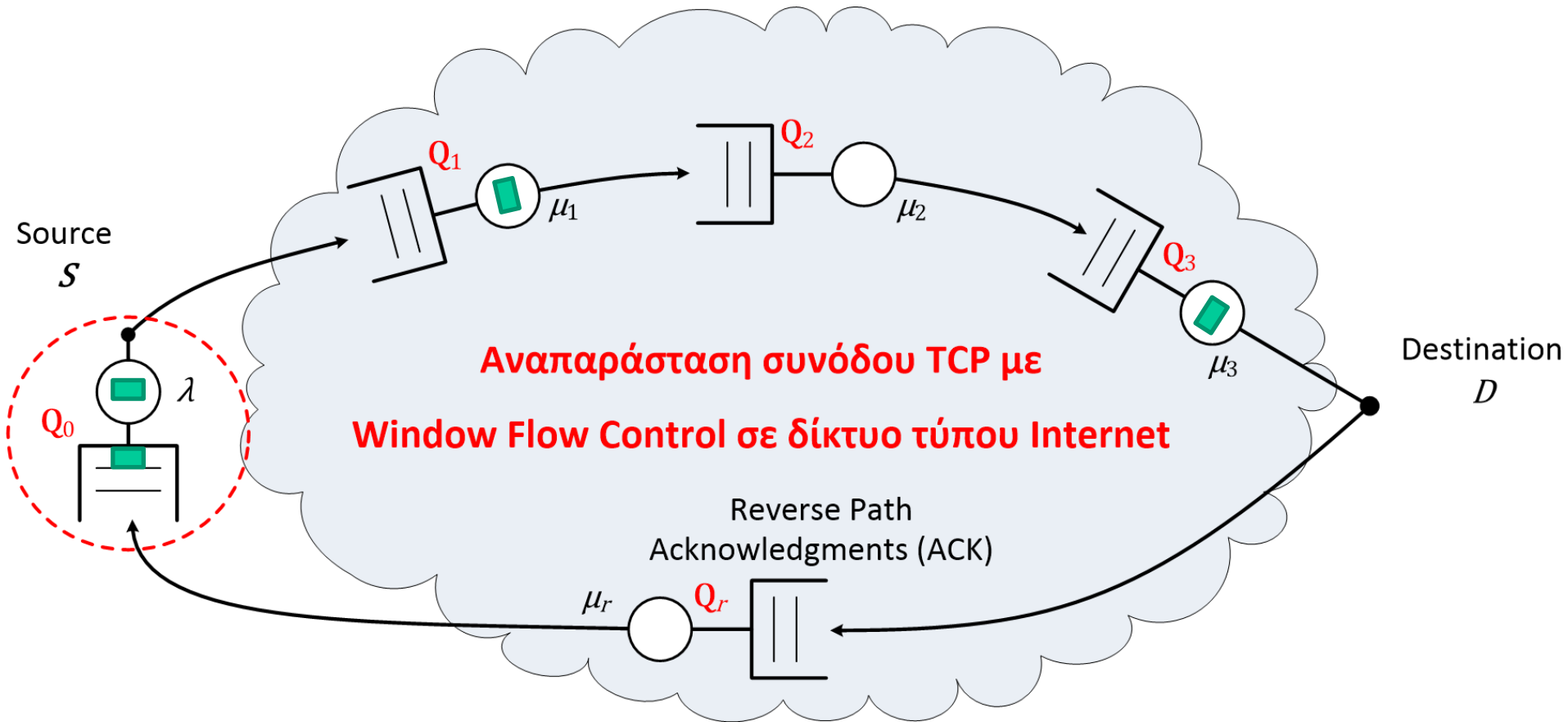
End-to-End Window Flow Control TCP, Μέγεθος Παραθύρου  $W$  (επανάληψη)





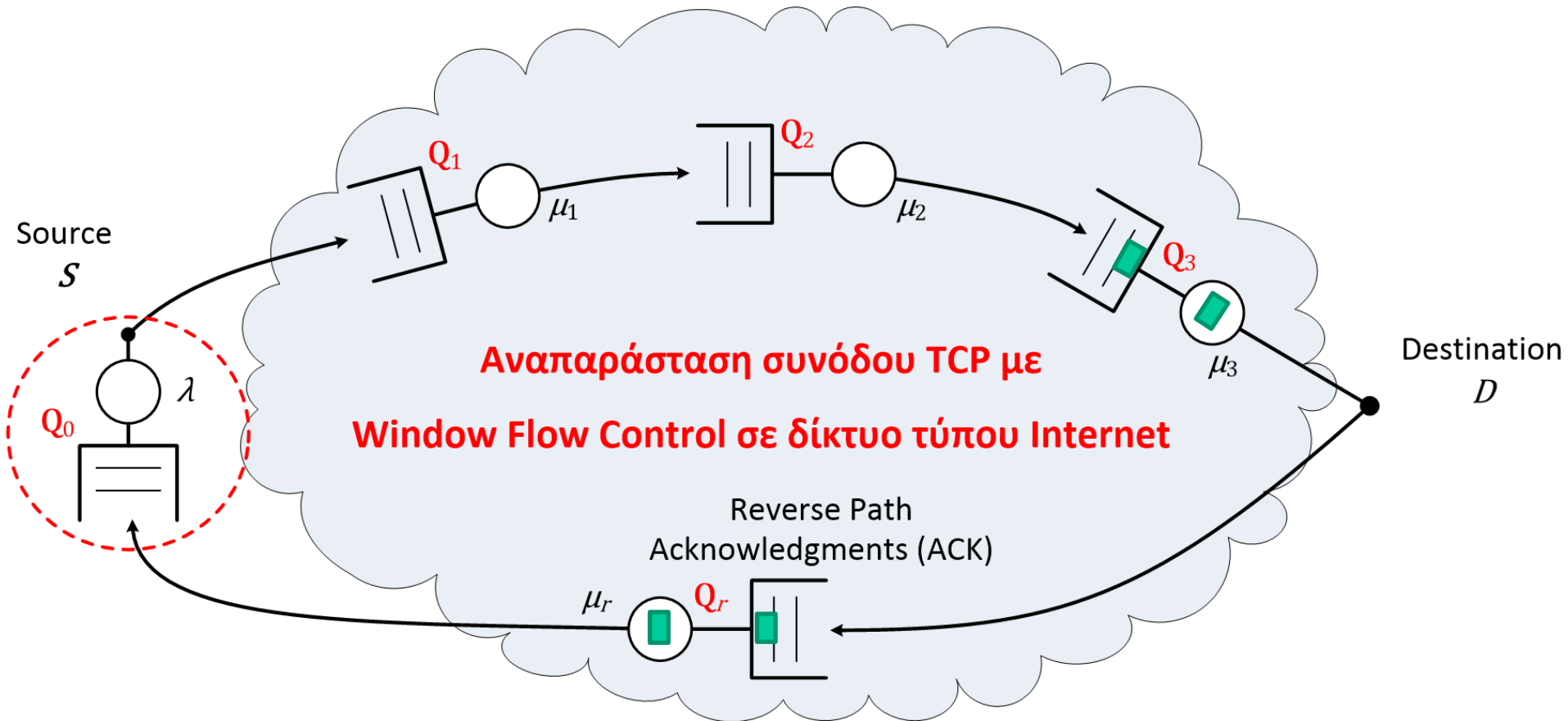
# ΕΛΕΓΧΟΣ ΡΟΗΣ ΑΠΟ ΑΚΡΟ ΣΕ ΑΚΡΟ ΣΤΟ INTERNET (4/10)

End-to-End Window Flow Control TCP, Μέγεθος Παραθύρου  $W$  (επανάληψη)



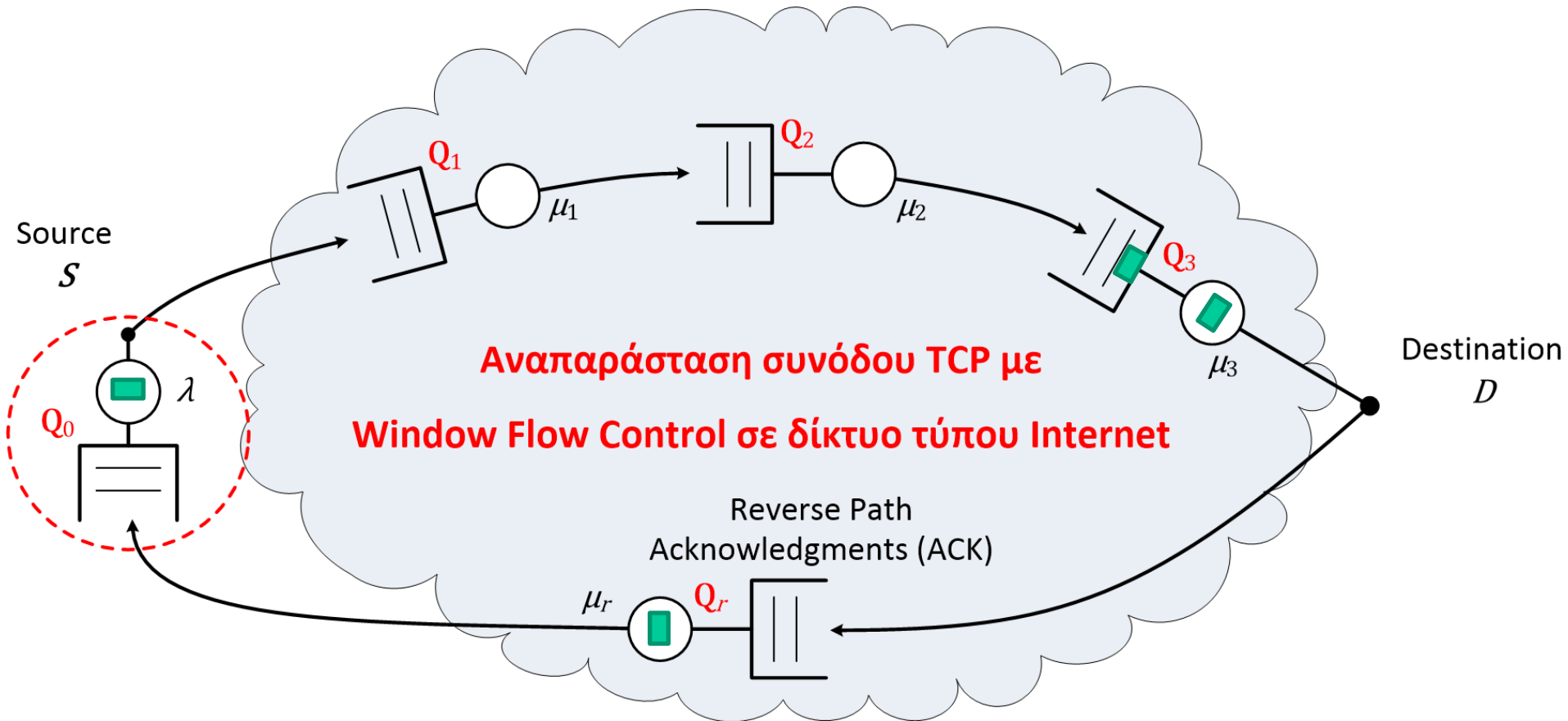
# ΕΛΕΓΧΟΣ ΡΟΗΣ ΑΠΟ ΑΚΡΟ ΣΕ ΑΚΡΟ ΣΤΟ INTERNET (5/10)

End-to-End Window Flow Control TCP, Μέγεθος Παραθύρου  $W$  (επανάληψη)



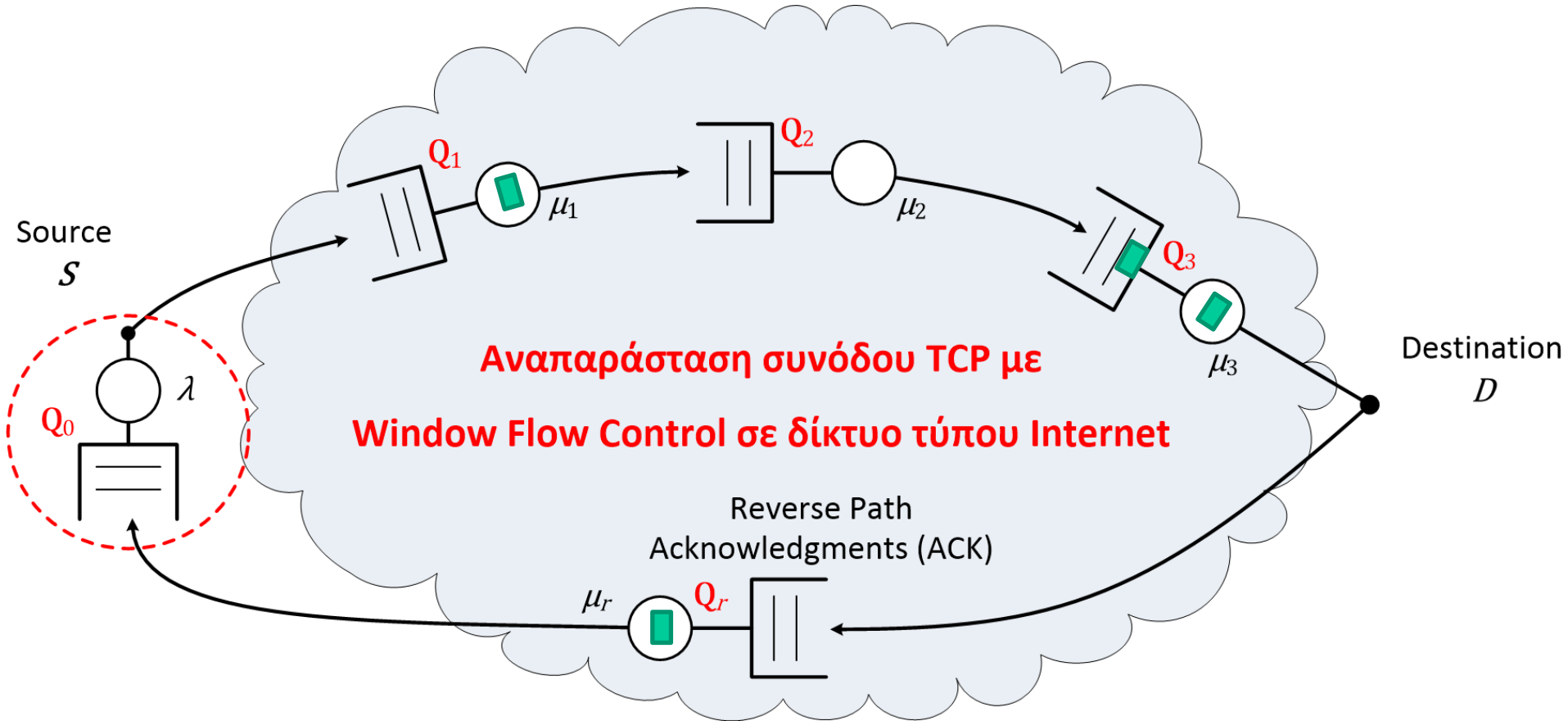
# ΕΛΕΓΧΟΣ ΡΟΗΣ ΑΠΟ ΑΚΡΟ ΣΕ ΑΚΡΟ ΣΤΟ INTERNET (6/10)

End-to-End Window Flow Control TCP, Μέγεθος Παραθύρου  $W$  (επανάληψη)



# ΕΛΕΓΧΟΣ ΡΟΗΣ ΑΠΟ ΑΚΡΟ ΣΕ ΑΚΡΟ ΣΤΟ INTERNET (7/10)

End-to-End Window Flow Control TCP, Μέγεθος Παραθύρου  $W$  (επανάληψη)



# ΕΛΕΓΧΟΣ ΡΟΗΣ ΑΠΟ ΑΚΡΟ ΣΕ ΑΚΡΟ ΣΤΟ INTERNET (8/10)

## End-to-End Window Flow Control TCP, Μέγεθος Παραθύρου $W$ (επανάληψη)

Υποθέτουμε πως  $\lambda = 1$ ,  $\mu_1 = \mu_2 = \mu_3 = 2$  πελάτες/sec  
 $\frac{1}{\mu_r} = \frac{1}{\mu_1} + \frac{1}{\mu_2} + \frac{1}{\mu_3} = \frac{3}{2}$  sec ή  $\mu_r = 2/3$  πελάτες/sec

Με  $X_0 = 1$  έχουμε:

$$\lambda X_0 = \mu_1 X_1 = \mu_2 X_2 = \mu_3 X_3 = \mu_r X_r$$

$$\text{Άρα: } X_1 = X_2 = X_3 = 0.5, X_r = 3/2$$

Ο αναδρομικός τύπος  $g(n, m) = g(n, m - 1) + X_m g(n - 1, m)$   
 δίνει τον πίνακα δεξιά

Οι σταθερές  $G(W) = g(W, 5)$  αντιστοιχούν σε  $W = 1, \dots, 8$

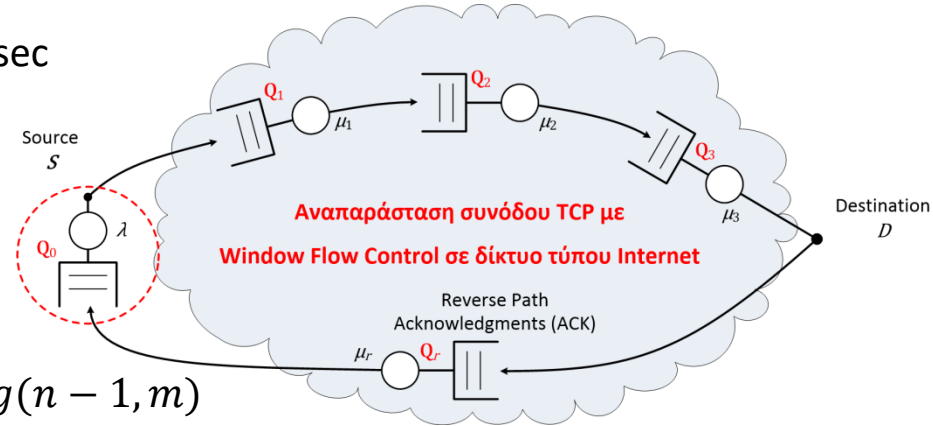
Η ρυθμαπόδοση του συστήματος  $\gamma$  σε πακέτα/sec είναι:

$$\gamma = \mu_1 p(n_1 \geq 1) = \mu_1 X_1 G(W - 1) / G(W)$$

Η μέση καθυστέρηση πακέτων σε sec από το S στο D είναι

$$E(T_{SD}) = [E(n_1) + E(n_2) + E(n_3)] / \gamma, \text{ όπου}$$

$$E[n_i] = \sum_{k=1}^N X_i^k \frac{G(N - k)}{G(N)}$$



Πίνακας Τιμών  $g(n, m)$

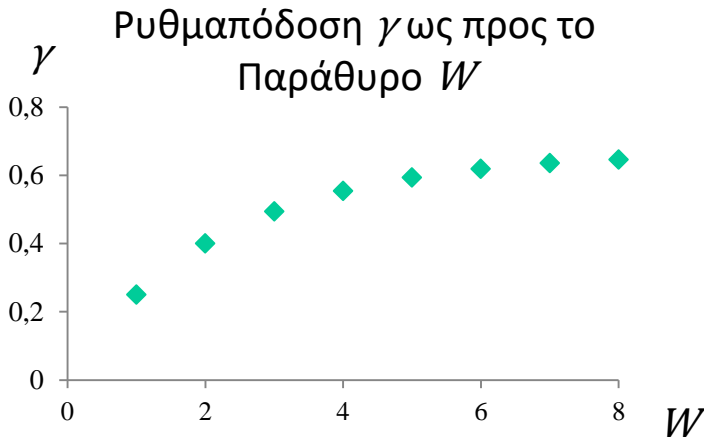
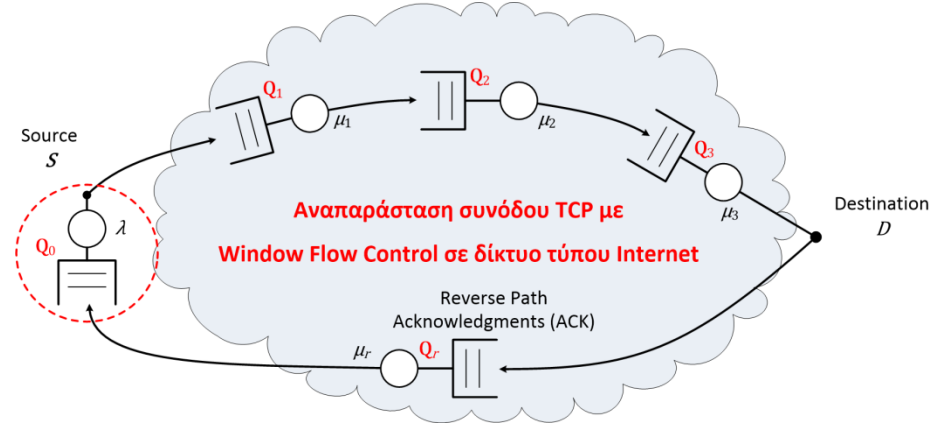
$W$	$X_0$	$X_1$	$X_2$	$X_3$	$X_r$
0	1	1	1	1	1
1	1	1.5	2	2.5	4
2	1	1.75	2.75	4	10
3	1	1.88	3.25	5.25	20.25
4	1	1.94	3.56	6.19	36.56
5	1	1.97	3.75	6.84	61.69
6	1	1.98	3.86	7.28	99.81
7	1	1.99	3.92	7.56	157.28
8	1	1.99	3.96	7.74	243.66

# ΕΛΕΓΧΟΣ ΡΟΗΣ ΑΠΟ ΑΚΡΟ ΣΕ ΑΚΡΟ ΣΤΟ INTERNET (9/10)

## End-to-End Window Flow Control TCP, Μέγεθος Παραθύρου $W$ (επανάληψη)

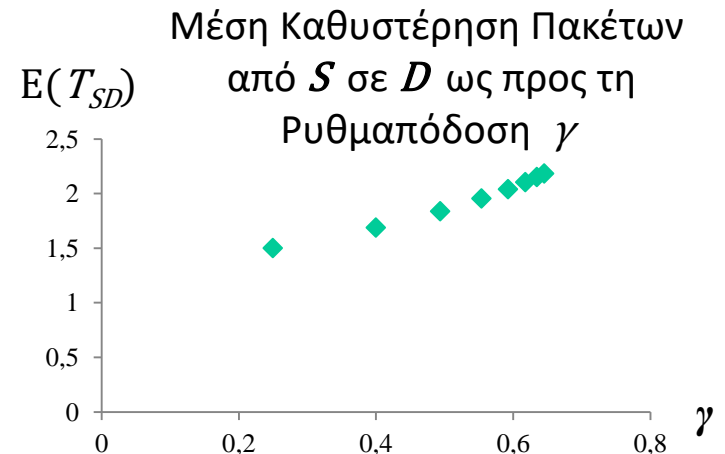
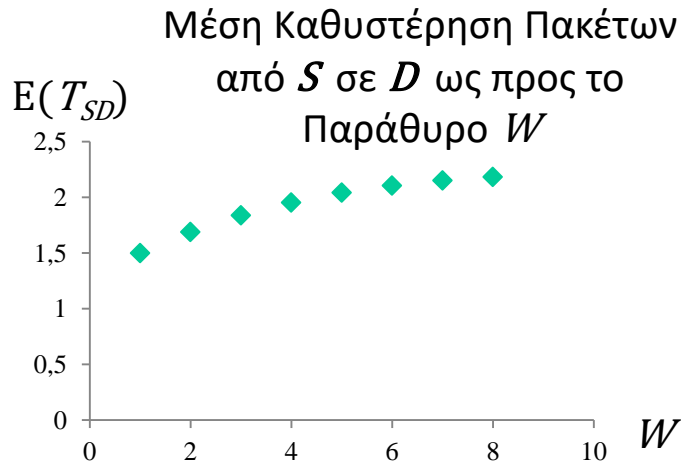
Ρυθμαπόδοση του συστήματος,  $\gamma = \mu_1 p(n_1 \geq 1)$ :

$W$	1	2	3	4	5	6	7	8
$\gamma$	0.25	0.4	0.49	0.55	0.59	0.62	0.63	0.65



Μέση καθυστέρηση πακέτου από το  $S$  στο  $D$ ,  $E(T_{SD})$ :

$W$	1	2	3	4	5	6	7	8
$E(T_{SD})$	1.5	1.69	1.84	1.95	2.04	2.1	2.15	2.18



# ΕΛΕΓΧΟΣ ΡΟΗΣ ΑΠΟ ΑΚΡΟ ΣΕ ΑΚΡΟ ΣΤΟ INTERNET (10/10)

## End-to-End Window Flow Control TCP, Μέγεθος Παραθύρου $W$ (επανάληψη)

### Σενάριο Συμφόρησης

Υποθέτουμε πως  $\lambda = 1$ ,  $\mu_1 = \mu_2 = \mu_3 = 1/2$  πελάτες/sec

$1/\mu_r = 1/\mu_1 + 1/\mu_2 + 1/\mu_3 = 6$  sec ή  $\mu_r = 1/6$

(ανεπαρκείς ταχύτητες γραμμών  $\Rightarrow$  **συμφόρηση**)

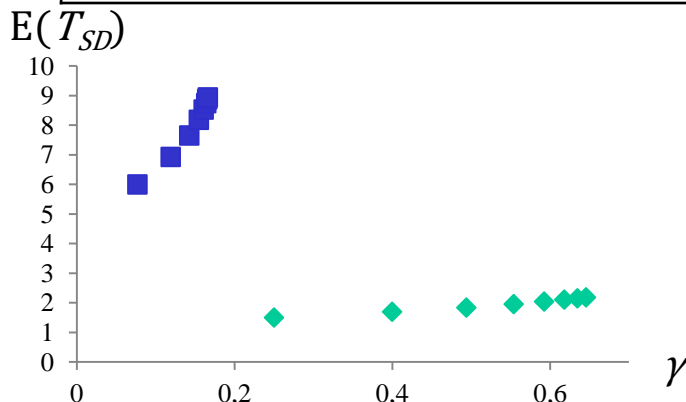
Έχουμε  $X_0 = 1$ ,  $X_1 = X_2 = X_3 = 2$ ,  $X_r = 6$

Η ρυθμαπόδοση του συστήματος  $\gamma$  είναι:

$W$	1	2	3	4	5	6	7	8
$\gamma$	0.077	0.119	0.142	0.155	0.161	0.164	0.165	0.166

Η μέση καθυστέρηση πακέτων  $E(T_{SD})$  είναι:

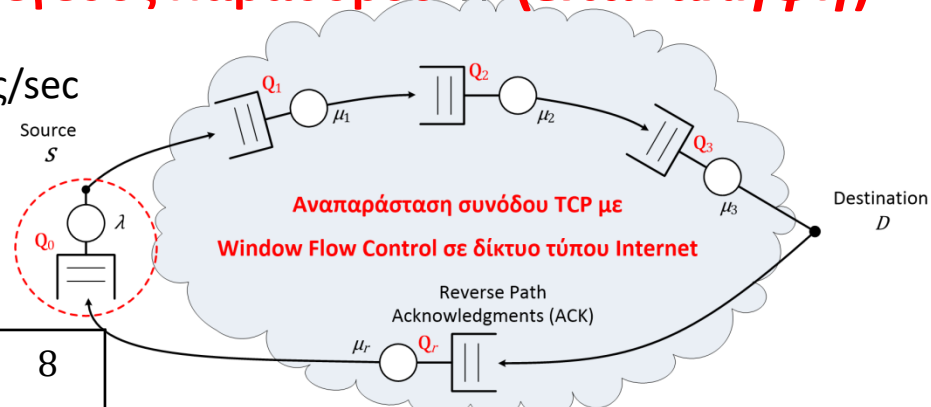
$W$	1	2	3	4	5	6	7	8
$E(T_{SD})$	6	6.92	7.65	8.18	8.53	8.75	8.87	8.94



### Σύγκριση Επίδοσης Σεναρίων για Αυξανόμενες τιμές του $W$

■ Συμφόρηση στο Δίκτυο ( $\mu_1 = \mu_2 = \mu_3 = 0.5, \lambda = 1$ )

◆ Δίκτυο με καλή επίδοση ( $\mu_1 = \mu_2 = \mu_3 = 2, \lambda = 1$ )



Πίνακας Τιμών  $g(n, m)$

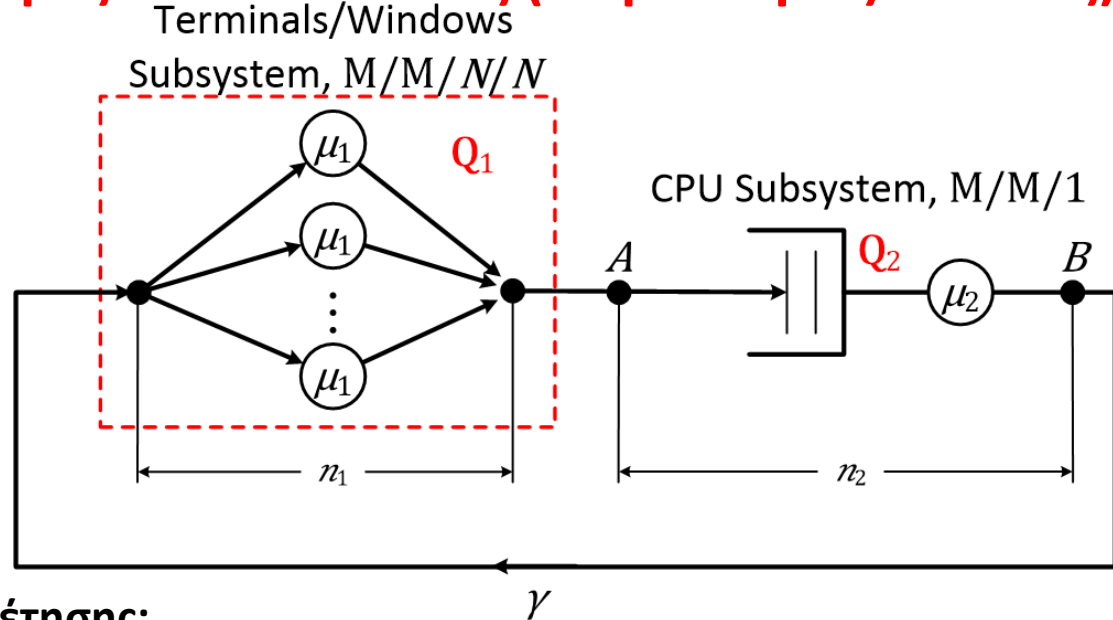
$W$	$X_0$	$X_1$	$X_2$	$X_3$	$X_r$
0	1	1	1	1	1
1	1	3	5	7	13
2	1	7	17	31	109
3	1	15	49	111	765
4	1	31	129	351	4941
5	1	63	321	1023	30669
6	1	127	769	2815	186829
7	1	255	1793	7423	1128397
8	1	511	4097	18943	6789325

**Στην περίπτωση συμφόρησης, ανεκτή καθυστέρηση απαιτεί μικρές τιμές του  $W$  με σημαντικούς περιορισμούς ωφέλιμης ρυθμαπόδοσης**

# ΠΑΡΑΔΕΙΓΜΑ ΠΟΛΥ-ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΥ (1/4)

## Κλειστό Δίκτυο με $M = 2$ Ουρές και $N$ Πελάτες (Παράλληλες Εντολές)

Το μοντέλο θεωρεί  $N$  πελάτες που περιφέρονται σε κλειστό δίκτυο. Εναλλάσσονται σε δύο υποσυστήματα, είτε με τη μορφή σκέψης ενός «χρήστη» (τερματικό ή ενεργό παράθυρο) για την παραγωγή (**input**) νέας εντολής, ή με τη μορφή επεξεργασίας εντολών που έχουν κατατεθεί και αναμένεται η απόκριση (**output**) στον «χρήστη»



### Μοντέλα 2 Υποσυστημάτων Εξυπηρέτησης:

- **Υποσύστημα Παραγωγής Εντολών:**  $Q_1$ ,  $M/M/\infty$  (ή  $M/M/N/N$ ) με εκθετικούς χρόνους εξυπηρέτησης (**Thinking Time**,  $T_{TH}$ ) μέσης τιμής  $E(T_{TH}) = 1/\mu_1$   
Ο χρόνος  $T_{TH}$  αντιστοιχεί με τον χρόνο «σκέψης» για την σύνταξη μιας νέας εντολής εισόδου (**input**) σε συνέχεια της απόκρισης (**output**) από προηγούμενη εντολή
- **Υποσύστημα Επεξεργασίας Εντολών:**  $Q_2$ ,  $M/M/1$  με εκθετικούς χρόνους εξυπηρέτησης  $1/\mu_2$  και μέσο χρόνο καθυστέρησης (**Processing Time**,  $T_{PR}$ ) μέσης τιμής  $E(T_{PR})$   
Η ουρά  $Q_2$  αντιστοιχεί με το συνολικό σύστημα επεξεργασίας εντολών (CPU, I/O) σαν συναθροισμένο ισοδύναμο μοντέλο (**aggregate equivalent**, γενικεύεται σαν **Ισοδύναμο Norton**)



# ΠΑΡΑΔΕΙΓΜΑ ΠΟΛΥ-ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΥ (2/4)

## Κλειστό Δίκτυο με $M = 2$ Ουρές και $N$ Πελάτες (Παράλληλες Εντολές)

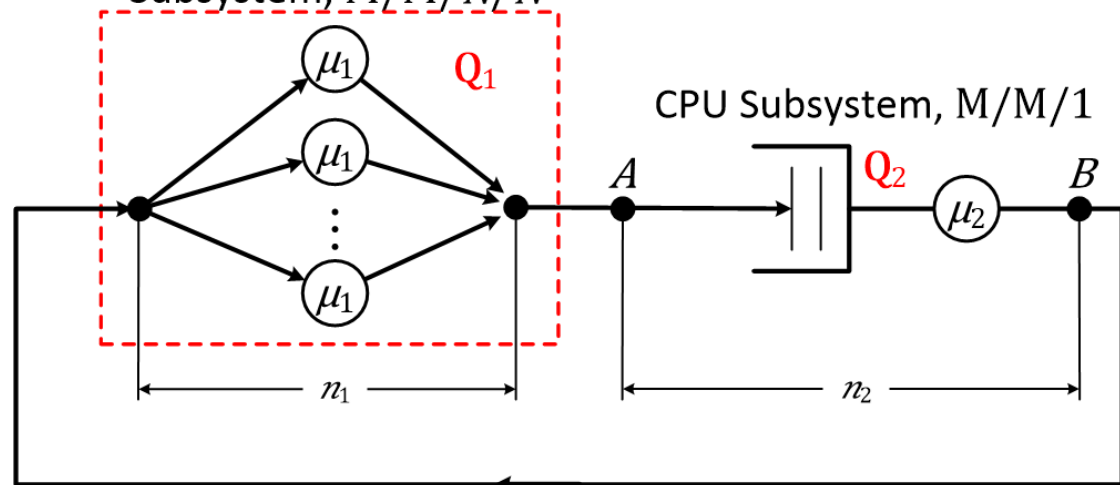
Terminals/Windows

Subsystem,  $M/M/N/N$

### Ανάλυση Κλειστού Δικτύου Μαρκοβ

Εργοδικές Καταστάσεις

$$\mathbf{n} = (n_1, n_2), n_1 + n_2 = N$$



### Εξισώσεις Ισοροπίας:

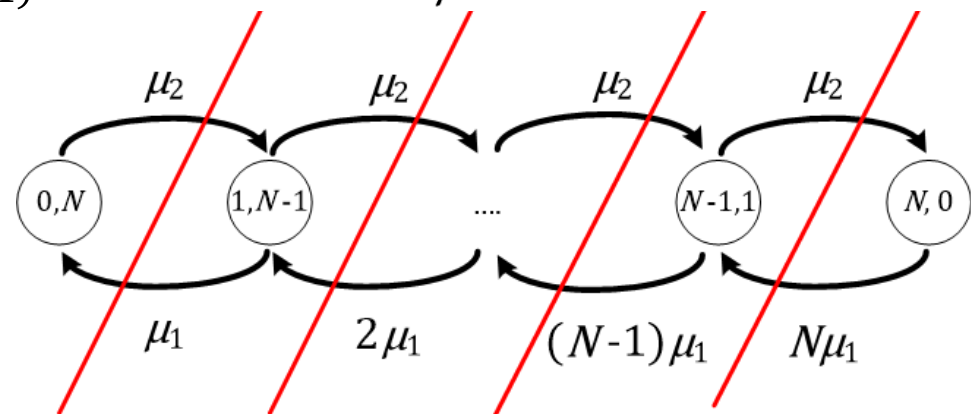
$$\mu_2 P(0, N) = \mu_1 P(1, N - 1)$$

$$\mu_2 P(k, N - k) = (k + 1) \mu_1 P(k + 1, N - k + 1)$$

$$\mu_2 P(N - 1, 1) = N \mu_1 P(N, 0)$$

Αν  $a \triangleq \mu_2 / \mu_1 \Rightarrow P(k, N - k) = P(0, N) \frac{a^k}{k!}$

και  $P(k, N - k) = \frac{\frac{a^k}{k!}}{\sum_{n=0}^N \frac{a^n}{n!}}$



Υπολογισμοί όπως για Erlang-B:

$$P(N, 0) = B(a, N), \quad P(N - 1, 1) = \frac{N \mu_1}{\mu_2} P(N, 0) = \frac{N}{a} B(a, N), \dots$$

$E(n_2) = \sum_{k=1}^N k P(N - k, k)$  και η ρυθμαπόδοση  $\gamma = \mu_2 [1 - P(N, 0)]$  εντολές/sec

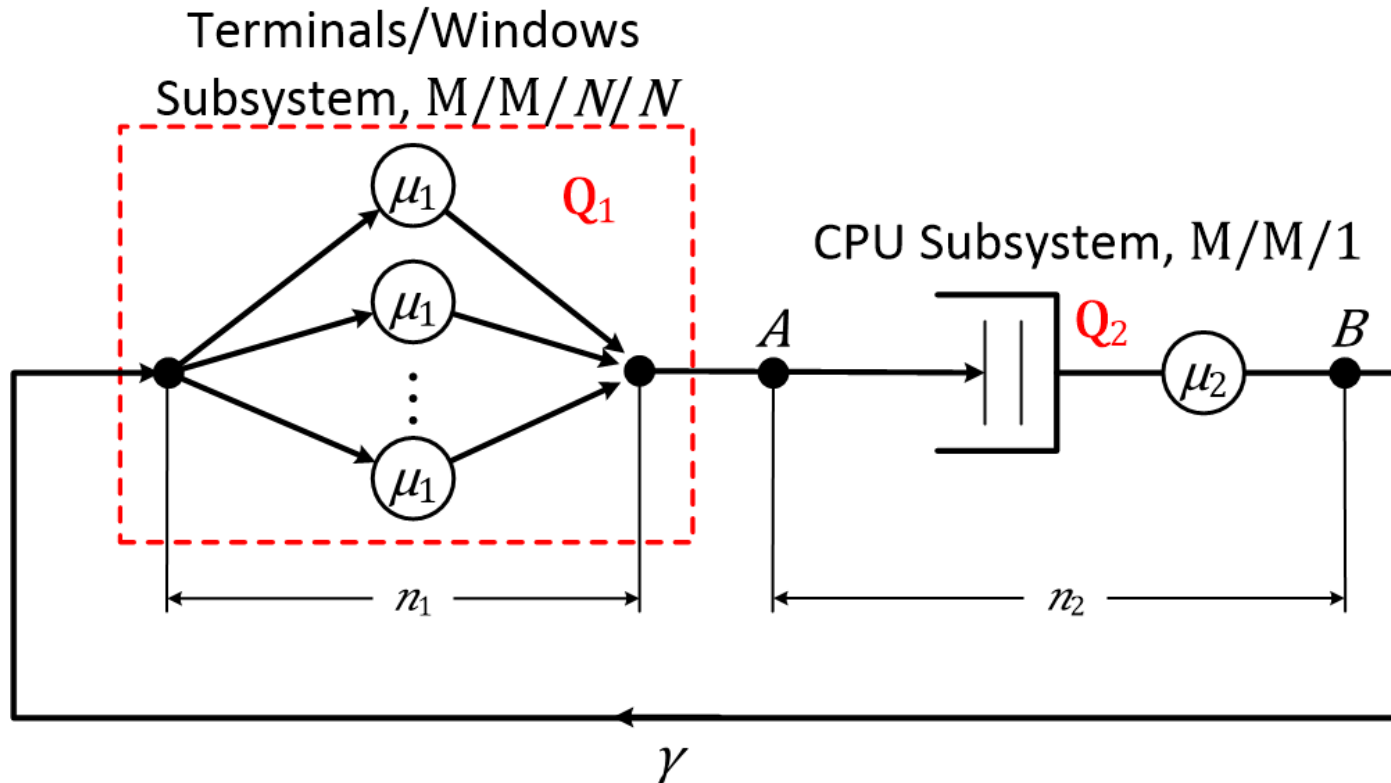
- Μέση Καθυστέρηση στο Υποσύστημα Επεξεργασίας:  $E(T_{PR}) = E(n_2) / \gamma$  sec
- Μέση Καθυστέρηση στο Υποσύστημα Τερματικών:  $E(T_{TH}) = 1 / \mu_1 = E(n_1) / \gamma = (N - E(n_2)) / \gamma$  sec

# ΠΑΡΑΔΕΙΓΜΑ ΠΟΛΥ-ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΥ (3/4)

## Κλειστό Δίκτυο με $M = 2$ Ουρές και $N$ Πελάτες (Παράλληλες Εντολές)

Αριθμητικά Αποτελέσματα:

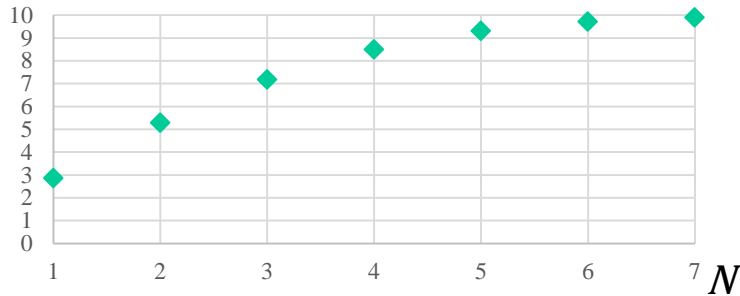
Εφαρμογή της ανάλυσης για τιμές των  $\mu_1 = 4, \mu_2 = 10$  ( $\alpha = \frac{\mu_2}{\mu_1} = 2.5$ ) και μεταβάλλοντας τις δυνατότητες παραλληλισμού (Βαθμός Πολύ-προγραμματισμού, **Degree of Parallelism - Multitasking**)  $N = 1, \dots, 7$



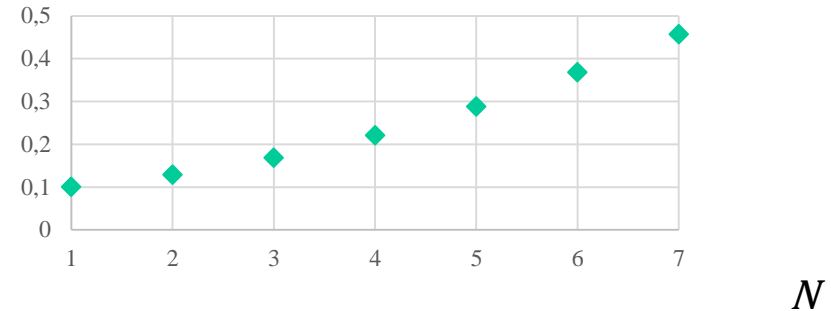
# ΠΑΡΑΔΕΙΓΜΑ ΠΟΛΥ-ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΥ (4/4)

Κλειστό Δίκτυο με  $M = 2$  Ουρές και  $N$  Πελάτες (Παράλληλες Εντολές)

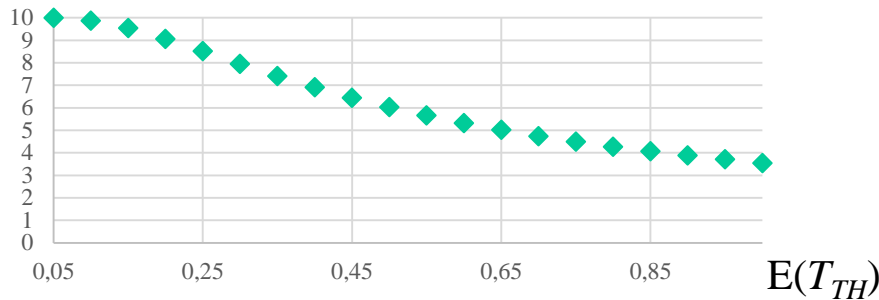
$\gamma$  ως προς  $N$ ,  $\mu_1=4$ ,  $\mu_2=10$



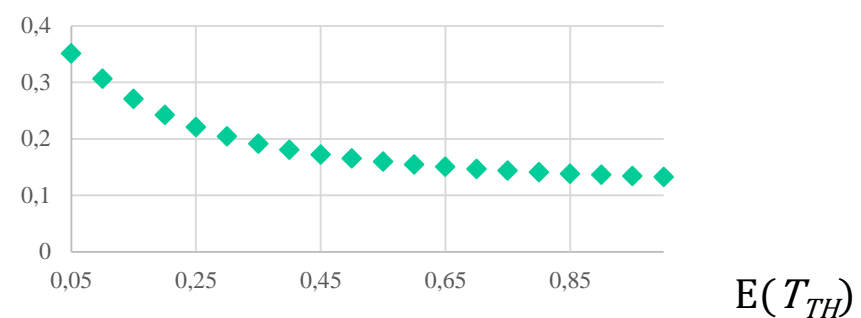
$E(T_{PR})$  ως προς  $N$ ,  $\mu_1=4$ ,  $\mu_2=10$



$\gamma$  ως προς  $E(T_{TH}) = 1/\mu_1$ ,  $\mu_2=10$ ,  $N=4$



$E(T_{PR})$  ως προς  $E(T_{TH}) = 1/\mu_1$ ,  $\mu_2=10$ ,  $N=4$



## Συμπεράσματα:

1. Για αυξανόμενο βαθμό πολύ-προγραμματισμού  $N$  παρατηρείται βελτίωση της Ρυθμαπόδοσης  $\gamma$  προς ένα μέγιστο όριο, με παράλληλη αύξηση της Μέσης Καθυστέρησης Επεξεργασίας Εντολών  $E(T_{PR})$
2. Για  $N = 4$  και  $\mu_2 = 10$  εντολές/sec όσο μειώνεται ο Μέσος Χρόνος Σκέψης  $E(T_{TH}) = 1/\mu_1 \rightarrow 0$  βελτιώνεται η Ρυθμαπόδοση  $\gamma \rightarrow \mu_2 = 10$  Εντολές/sec και η Μέση Καθυστέρηση Επεξεργασίας Εντολών  $E(T_{PR}) \rightarrow 0.36$  sec

# ΑΝΑΛΥΣΗ ΜΕΣΗΣ ΤΙΜΗΣ ΓΙΑ ΚΛΕΙΣΤΑ ΔΙΚΤΥΑ ΕΚΘΕΤΙΚΩΝ ΟΥΡΩΝ

## MEAN VALUE ANALYSIS - MVA (1/5)

1. *Mean Value Analysis of Closed Multichain Queuing Networks*, M. Reiser & S.S. Lavenberg, Journal of the ACM 27 (2), 1980 (<https://dl.acm.org/citation.cfm?doid=322186.322195>)
2. *Wikipedia* [https://en.wikipedia.org/wiki/Mean\\_value\\_analysis](https://en.wikipedia.org/wiki/Mean_value_analysis)

### Παραδοχές:

- Κλειστό δίκτυο από  $M$  διασυνδεδεμένα υποσυστήματα (ουρές)  $Q_i, i = 1, 2, \dots, M$  στην εργοδική κατάσταση, έκαστο με  $n_i(k)$  πελάτες και ρυθμαπόδοση συστήματος  $\gamma(k)$  πελάτες/sec ανά περιφορά πελάτη, όταν περιφέρονται στο δίκτυο  $k = 1, 2, \dots, N$  πελάτες
- Ανεξάρτητοι εκθετικοί εξυπηρετητές  $i = 1, 2, \dots, M$  μέσου ρυθμού  $\mu_i$  με παραδοχή **Kleinrock**
- Τυχαία Δρομολόγηση από υποσύστημα  $Q_i$  σε  $Q_j$  με πιθανότητα  $p_{ij} = \text{Prob}\{i \rightarrow j\}$ . Η δρομολόγηση ορίζει τους μέσους ρυθμούς πελατών (ρυθμαποδόσεις)  $\lambda_i(k)$  στα υποσυστήματα  $Q_i$  σαν το γινόμενο της συνολικής ρυθμαπόδοσης  $\gamma(k)$  και του μέσου αριθμού επισκέψεων  $v_i$  ενός πελάτη στο  $Q_i$  ανά περιφορά του στο συνολικό σύστημα :  
$$\lambda_i(k) = v_i \cdot \gamma(k)$$

### Θεώρημα Αφίξεων σε Κλειστά Δίκτυα Εκθετικών Ουρών:

Πελάτης που αφικνείται σε υποσύστημα εξυπηρέτησης  $Q_i$  κλειστού δικτύου ανεξαρτήτων εκθετικών εξυπηρετήσεων συνολικού αριθμού πελατών  $N$  παρατηρεί την εργοδική κατάσταση ισορροπίας του δικτύου όπως προκύπτει με  $N - 1$  πελάτες (**βλέπει το σύστημα όπως θα ισορροπούσε χωρίς τη παρουσία του**)

# ΑΝΑΛΥΣΗ ΜΕΣΗΣ ΤΙΜΗΣ ΓΙΑ ΚΛΕΙΣΤΑ ΔΙΚΤΥΑ ΕΚΘΕΤΙΚΩΝ ΟΥΡΩΝ

## MEAN VALUE ANALYSIS - MVA (2/5)

### Αναδρομικός Αλγόριθμος:

Έστω:

- $E[n_i(k)]$ : Μέσος όρος πελατών στο  $Q_i$  με  $k = 1, 2, \dots, N$  πελάτες στο δίκτυο

Έχουμε

- Μέση καθυστέρηση πελατών στο  $Q_i$ :  $E[T_i(k)]$

$$E[n_i(k)] = E[T_i(k)] \cdot \lambda_i(k) \quad (\text{τύπος } \mathbf{Little} \text{ ανά υποσύστημα } Q_i)$$

- Συνολική μέση καθυστέρηση περιφοράς πελάτη  $\sum_{i=1}^M v_i \cdot E[T_i(k)]$ :

- Συνολική ρυθμαπόδοση  $\gamma(k) = \frac{k}{\sum_{i=1}^M v_i E[T_i(k)]}$  (τύπος **Little** για το δίκτυο με  $k$  πελάτες)

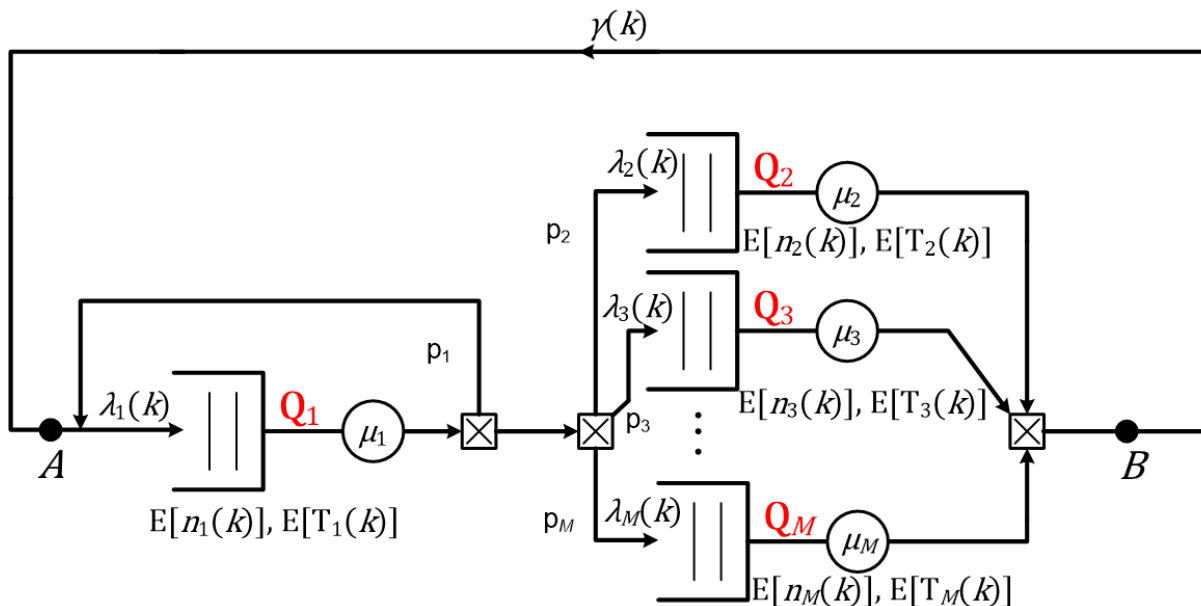
Από το **Θεώρημα Αφίξεων σε Κλειστά Δίκτυα** έχουμε τον αναδρομικό τύπο για  $k = 1, 2, \dots, N$

$$E[T_i(k)] = \frac{1 + E[n_i(k-1)]}{\mu_i}$$

με αρχικές τιμές  $E[n_i(0)] = 0, i = 1, 2, \dots, M$

# ΑΝΑΛΥΣΗ ΜΕΣΗΣ ΤΙΜΗΣ ΓΙΑ ΚΛΕΙΣΤΑ ΔΙΚΤΥΑ ΕΚΘΕΤΙΚΩΝ ΟΥΡΩΝ

## MEAN VALUE ANALYSIS - MVA (3/5)



Εφαρμογή για:  $N = 3$  πελάτες,  $M = 3$  ουρές και τιμές των παραμέτρων:

$$\mu_1 = \frac{1}{28} \text{msec}^{-1}, \mu_2 = \frac{1}{40} \text{msec}^{-1}, \mu_3 = \frac{1}{280} \text{msec}^{-1} \text{ και } p_1 = 0.1, p_2 = 0.7, p_3 = 0.2$$

Εύρεση των  $v_i$ : Σε κάθε περιφορά πελάτη στον εξωτερικό βρόχο ανάδρασης αντιστοιχούν  $v_i$  επισκέψεις στα υποσυστήματα  $Q_i$  που συντελούν στις ρυθμαποδόσεις  $\lambda_i(k)$ :

$$\lambda_i(k) = v_i \cdot \gamma(k) \text{ ή } \lambda_1(k) = \gamma(k) + p_1 \cdot \lambda_1(k), \lambda_2(k) = p_2 \cdot \lambda_1(k), \lambda_3(k) = p_3 \cdot \lambda_1(k) \Rightarrow$$

$$v_1 = \frac{\lambda_1(k)}{\gamma(k)} = \frac{\lambda_1(k)}{\lambda_1(k)[1-p_1]} = \frac{10}{9}, v_2 = p_2 v_1 = \frac{7}{9}, v_3 = p_3 v_1 = \frac{2}{9}$$

# ΑΝΑΛΥΣΗ ΜΕΣΗΣ ΤΙΜΗΣ ΓΙΑ ΚΛΕΙΣΤΑ ΔΙΚΤΥΑ ΕΚΘΕΤΙΚΩΝ ΟΥΡΩΝ

## MEAN VALUE ANALYSIS - MVA (4/5)

Για  $k = 1$ :

$$E[T_1(1)] = \frac{1+E[n_1(0)]}{\mu_1} = \frac{1+0}{1/28} = 28 \text{ msec} = 0.028 \text{ sec}$$

$$E[T_2(1)] = \frac{1+E[n_2(0)]}{\mu_2} = \frac{1+0}{1/40} = 40 \text{ msec} = 0.040 \text{ sec}$$

$$E[T_3(1)] = \frac{1+E[n_3(0)]}{\mu_3} = \frac{1+0}{1/280} = 280 \text{ msec} = 0.280 \text{ sec}$$

$$\gamma(1) = \frac{1}{v_1 E[T_1(1)] + v_2 E[T_2(1)] + v_3 E[T_3(1)]} = \frac{1}{\frac{10}{9} \cdot 0.028 + \frac{7}{9} \cdot 0.040 + \frac{2}{9} \cdot 0.280} = 8.03572 \text{ πελάτες/sec}$$

$$\lambda_1(1) = v_1 \cdot \gamma(1) = \frac{10}{9} \cdot 8.03572 = 8.92858, \lambda_2(1) = v_2 \cdot \gamma(1) = \frac{7}{9} \cdot 8.03572 = 6.25000,$$

$$\lambda_3(1) = v_3 \cdot \gamma(1) = \frac{2}{9} \cdot 8.03572 = 1.78572 \text{ (πελάτες/sec)}$$

$$E[n_1(1)] = E[T_1(1)] \cdot \lambda_1(1) = 0.028 \cdot 8.92858 = 0.25 \text{ πελάτες}$$

$$E[n_2(1)] = E[T_2(1)] \cdot \lambda_2(1) = 0.040 \cdot 6.25000 = 0.25 \text{ πελάτες}$$

$$E[n_3(1)] = E[T_3(1)] \cdot \lambda_3(1) = 0.280 \cdot 1.78572 = 0.5 \text{ πελάτες}$$

Για  $k = 2$ :

$$E[T_1(2)] = \frac{1+E[n_1(1)]}{\mu_1} = \frac{1+0.25}{1/28} = 35 \text{ msec} = 0.035 \text{ sec}$$

$$E[T_2(2)] = \frac{1+E[n_2(1)]}{\mu_2} = \frac{1+0.25}{1/40} = 50 \text{ msec} = 0.050 \text{ sec}$$

$$E[T_3(2)] = \frac{1+E[n_3(1)]}{\mu_3} = \frac{1+0.5}{1/280} = 420 \text{ msec} = 0.420 \text{ sec}$$

$$\gamma(2) = \frac{1}{v_1 E[T_1(2)] + v_2 E[T_2(2)] + v_3 E[T_3(2)]} = \frac{1}{\frac{10}{9} \cdot 0.035 + \frac{7}{9} \cdot 0.050 + \frac{2}{9} \cdot 0.420} = 11.68832 \text{ πελάτες/sec}$$

$$\lambda_1(2) = v_1 \cdot \gamma(2) = \frac{10}{9} \cdot 11.68832 = 12.98702, \lambda_2(2) = v_2 \cdot \gamma(2) = \frac{7}{9} \cdot 11.68832 = 9.09092,$$

$$\lambda_3(2) = v_3 \cdot \gamma(2) = \frac{2}{9} \cdot 11.68832 = 2.59740 \text{ (πελάτες/sec)}$$

$$E[n_1(2)] = E[T_1(2)] \cdot \lambda_1(2) = 0.035 \cdot 12.98702 = 0.45455 \text{ πελάτες}$$

$$E[n_2(2)] = E[T_2(2)] \cdot \lambda_2(2) = 0.050 \cdot 9.09092 = 0.45455 \text{ πελάτες}$$

$$E[n_3(2)] = E[T_3(2)] \cdot \lambda_3(2) = 0.420 \cdot 2.59740 = 1.09090 \text{ πελάτες}$$

# ΑΝΑΛΥΣΗ ΜΕΣΗΣ ΤΙΜΗΣ ΓΙΑ ΚΛΕΙΣΤΑ ΔΙΚΤΥΑ ΕΚΘΕΤΙΚΩΝ ΟΥΡΩΝ

## MEAN VALUE ANALYSIS - MVA (5/5)

Για  $k = 3$ :

$$E[T_1(3)] = \frac{1+E[n_1(2)]}{\mu_1} = \frac{1+0.45455}{1/28} = 40.72 \text{ msec} = 0.0407 \text{ sec}$$

$$E[T_2(3)] = \frac{1+E[n_2(2)]}{\mu_2} = \frac{1+0.45455}{1/40} = 58.18 \text{ msec} = 0.058 \text{ sec}$$

$$E[T_3(3)] = \frac{1+E[n_3(2)]}{\mu_3} = \frac{1+1.09090}{1/280} = 585.45 \text{ msec} = 0.585 \text{ sec}$$

$$\gamma(3) = \frac{1}{v_1 E[T_1(3)] + v_2 E[T_2(3)] + v_3 E[T_3(3)]} = \frac{3}{\frac{10}{9} \cdot 0.0407 + \frac{7}{9} \cdot 0.058 + \frac{2}{9} \cdot 0.585} = 13.616 \text{ πελάτες/sec}$$

$$\lambda_1(3) = v_1 \cdot \gamma(3) = \frac{10}{9} \cdot 13.616 = 15.1289, \lambda_2(3) = v_2 \cdot \gamma(3) = \frac{7}{9} \cdot 13.616 = 10.5902,$$

$$\lambda_3(3) = v_3 \cdot \gamma(3) = \frac{2}{9} \cdot 13.616 = 3.0258 \text{ (πελάτες/sec)}$$

$$E[n_1(3)] = E[T_1(3)] \cdot \lambda_1(3) = 0.0407 \cdot 15.1289 = 0.61575 \text{ πελάτες}$$

$$E[n_2(3)] = E[T_2(3)] \cdot \lambda_2(3) = 0.058 \cdot 10.5902 = 0.61423 \text{ πελάτες}$$

$$E[n_3(3)] = E[T_3(3)] \cdot \lambda_3(3) = 0.585 \cdot 3.0258 = 1.77001 \text{ πελάτες}$$

Για  $k = 4$ :

$$E[T_1(4)] = \frac{1+E[n_1(3)]}{\mu_1} = \frac{1+0.61575}{1/28} = 45.24 \text{ msec} = 0.045 \text{ sec}$$

$$E[T_2(4)] = \frac{1+E[n_2(3)]}{\mu_2} = \frac{1+0.61423}{1/40} = 64.57 \text{ msec} = 0.065 \text{ sec}$$

$$E[T_3(4)] = \frac{1+E[n_3(3)]}{\mu_3} = \frac{1+0.585}{1/280} = 443.8 \text{ msec} = 0.444 \text{ sec}$$

$$\gamma(4) = \frac{1}{v_1 E[T_1(4)] + v_2 E[T_2(4)] + v_3 E[T_3(4)]} = \frac{4}{\frac{10}{9} \cdot 0.045 + \frac{7}{9} \cdot 0.065 + \frac{2}{9} \cdot 0.444} = 20.078 \text{ πελάτες/sec}$$

$$\lambda_1(4) = v_1 \cdot \gamma(4) = \frac{10}{9} \cdot 20.078 = 22.3089, \lambda_2(4) = v_2 \cdot \gamma(4) = \frac{7}{9} \cdot 20.078 = 15.6162,$$

$$\lambda_3(4) = v_3 \cdot \gamma(4) = \frac{2}{9} \cdot 20.078 = 4.4618 \text{ (πελάτες/sec)}$$

$$E[n_1(4)] = E[T_1(4)] \cdot \lambda_1(4) = 0.045 \cdot 22.3089 = 1.00390 \text{ πελάτες}$$

$$E[n_2(4)] = E[T_2(4)] \cdot \lambda_2(4) = 0.065 \cdot 15.6162 = 1.01505 \text{ πελάτες}$$

$$E[n_3(4)] = E[T_3(4)] \cdot \lambda_3(4) = 0.444 \cdot 4.4618 = 1.98104 \text{ πελάτες}$$



# ΠΡΟΣΟΜΟΙΩΣΗ ΣΤΟΧΑΣΤΙΚΟΥ ΣΥΣΤΗΜΑΤΟΣ MARKOV

## (**BIRTH-DEATH**)

- Έστω στοχαστικό σύστημα **Birth-Death** με απ' ευθείας μεταβάσεις από την εργοδική κατάσταση  $S_i$  σε γειτονική κατάσταση  $S_j$  που μπορεί να γίνονται με ρυθμό  $r_{ij}$

- Ο συνολικός ρυθμός μεταβάσεων από την  $S_i$  προς όλες τις γειτονικές καταστάσεις είναι

$$R_i = \sum_{S_j} r_{ij}$$

- Ο χρόνος παραμονής (**dwelt time**)  $T_i$  στη κατάσταση  $S_i$  μέχρι την επόμενη μετάβαση είναι τυχαία μεταβλητή που προκύπτει από το **ελάχιστο** ανεξαρτήτων εκθετικών τυχαίων μεταβλητών, αντίστοιχων με όλες τις πιθανές μεταβάσεις εκτός της  $S_i$ . Η  $T_i$  θα είναι εκθετική με ρυθμό  $R_i = \sum_{S_j} r_{ij}$  και μέση τιμή

$$E[T_i] = \frac{1}{R_i}$$

- Οι εργοδικές πιθανότητες  $P[S_i]$  αν υπάρχουν μπορούν να υπολογισθούν σαν το όριο του λόγου του συνολικού χρόνου που το σύστημα παραμένει στη κατάσταση  $S_i$  δια του χρονικού διαστήματος παρατήρησης  $T$  μιας χρονικής εξέλιξης (δείγματος) της στοχαστικής ανέλιξης :

$$P[S_i] = \lim_{T \rightarrow \infty} \frac{\#\{\text{ΕΠΙΣΚΕΨΕΩΝ στην } S_i\} \times E[T_i]}{T} \approx \frac{\#\{\text{ΕΠΙΣΚΕΨΕΩΝ στην } S_i\} \times E[T_i]}{\sum_{S_j} \{\#\{\text{ΕΠΙΣΚΕΨΕΩΝ στην } S_j\} \times E[T_j]\}}$$

**Άρα μπορούμε να προσομοιώσουμε σύστημα Birth-Death καταμετρώντας τις αφίξεις στις διάφορες καταστάσεις που μεταβαίνει**

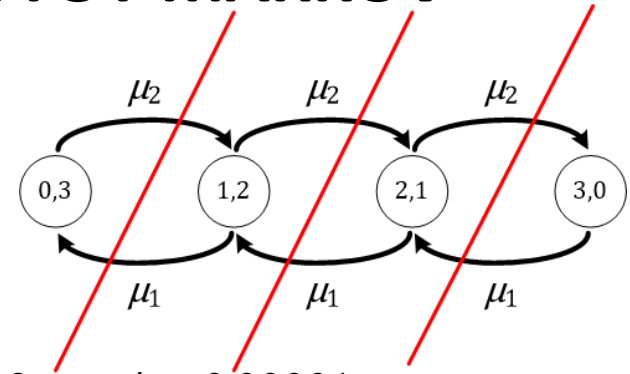
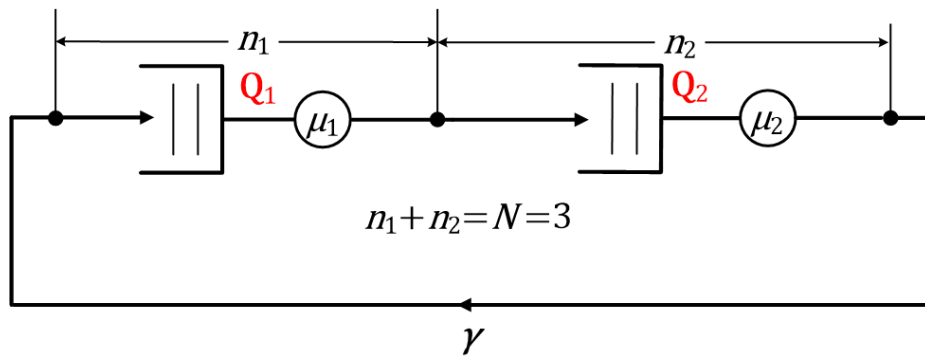
- Η εξέλιξη της κατάστασης (πληθυσμού) του συστήματος προκύπτει από τις πιθανότητες μετάβασης από την κατάσταση  $S_i$  στις γειτονικές καταστάσεις  $S_j$  με το δεδομένο ότι μια από αυτές θα συμβεί με απόλυτη βεβαιότητα:

$$P[S_i \rightarrow S_j / \text{μετάβαση}] = \frac{r_{ij}}{R_i}$$

- Η προσομοίωση ενεργοποιεί τις μεταβάσεις με κλήση τυχαίου αριθμού  $RANDOM(0,1)$  ομοιόμορφα κατανεμημένου μεταξύ  $(0, 1)$ :

Αν ο  $RANDOM(0,1)$  ανήκει σε υποδιάστημα του  $(0, 1)$  **ανάλογο** του  $\frac{r_{ij}}{R_i}$ , η επόμενη κατάσταση θα είναι η  $S_j$

# ΠΡΟΣΟΜΟΙΩΣΗ ΚΛΕΙΣΤΟΥ ΔΙΚΤΥΟΥ ΜΑΡΚΟΒ



**Κριτήριο σύγκλισης:** Διαδοχικές τιμές μέσου αριθμού πελατών σε κάθε ουρά  $< 0.00001$

- $\mu_1/\mu_2 = 1$

$P(0,3) = 0.252, P(1,2) = 0.251, P(2,1) = 0.249, P(3,0) = 0.247, E(n_1) = 1.49, E(n_2) = 1.51, E(n_1) + E(n_2) = 3$

- $\mu_1/\mu_2 = 2$

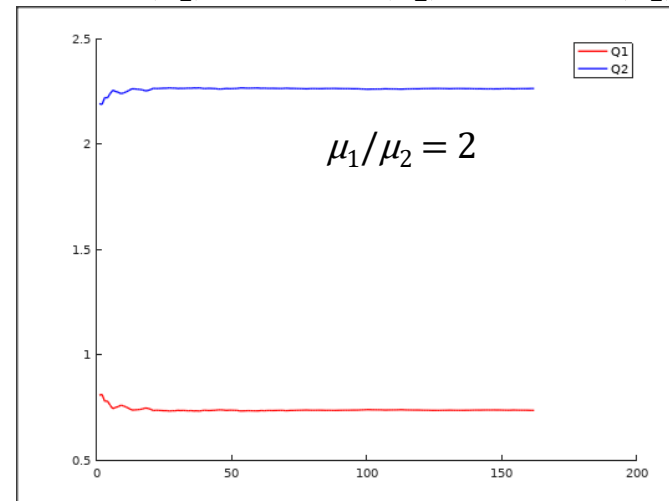
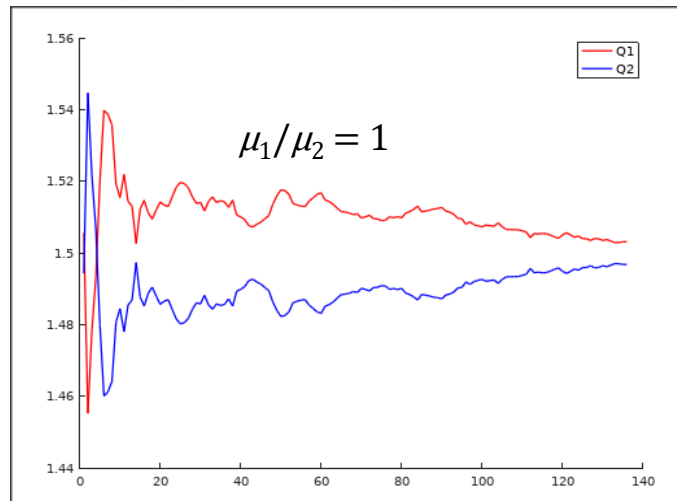
$P(0,3) = 0.534, P(1,2) = 0.266, P(2,1) = 0.133, P(3,0) = 0.067, E(n_1) = 0.73, E(n_2) = 2.27, E(n_1) + E(n_2) = 3$

- $\mu_1/\mu_2 = 3$

$P(0,3) = 0.673, P(1,2) = 0.225, P(2,1) = 0.249, P(3,0) = 0.025, E(n_1) = 0.45, E(n_2) = 2.55, E(n_1) + E(n_2) = 3$

- $\mu_1/\mu_2 = 4$

$P(0,3) = 0.754, P(1,2) = 0.158, P(2,1) = 0.047, P(3,0) = 0.012, E(n_1) = 0.32, E(n_2) = 2.68, E(n_1) + E(n_2) = 3$



Διαγράμματα μέσου αριθμού πελατών  $E(n_1), E(n_2)$  στα υποσυστήματα  $Q_1, Q_2$

# ΓΕΝΙΚΕΥΣΗ ΓΙΑ ΔΙΚΤΥΑ ΟΥΡΩΝ ΜΕ ΕΡΓΟΔΙΚΕΣ ΠΙΘΑΝΟΤΗΤΕΣ ΜΟΡΦΗΣ ΓΙΝΟΜΕΝΟΥ

- **BCMP Networks:** F. **B**asket, K.M. **C**handi, R.H. **M**untz, F.C. **P**alacios: “*Open, Closed and Mixed Networks of Queues with Different Classes of Customers*”, Journal of the ACM, 22 (2), April 1975
- **Παραδοχές:**
  - Δίκτυο  $M$  συστημάτων εξυπηρέτησης (ουρών)  $Q_i$  των εξής τύπων:
    1. Εξυπηρέτησης **FCFS (FIFO)  $M/M/1$**  με εκθετικό εξυπηρετητή  $1/\mu_i$  και ενιαία **μοναδική κλάση (*chain*)** πελατών
    2. Εξυπηρέτησης **Processor Sharing  $M/G/1$**  με πολλαπλές κλάσεις (***chains***) πελατών
    3. Ουρές με άπειρους εξυπηρετητές  **$M/G/\infty$**  με πολλαπλές κλάσεις (***chains***) πελατών
    4. Εξυπηρέτησης **LCFS (with pre-emptive resume)  $M/G/1$**  με πολλαπλές κλάσεις (***chains***) πελατών
  - Για τις περιπτώσεις 2-4 η κατανομή του χρόνου εξυπηρέτησης πρέπει να έχει μετασχηματισμό **Laplace** μορφής κλάσματος (rational Laplace Transform)
  - Η δρομολόγηση κλάσεων μεταξύ ουρών γίνεται με τυχαίο τρόπο
  - Ισχύει η παραδοχή ανεξαρτησίας του **Kleinrock**
  - Poisson εξωτερικές αφίξεις
- **Αποτέλεσμα:** Η εργοδική πιθανότητα (αν υπάρχει) του διανύσματος κατάστασης  $(x_1, x_2, \dots, x_M)$  του δικτύου δίνεται σε μορφή γινομένου παραγόντων εξαρτώμενων από την κατάσταση της κάθε ουράς:  $p(x_1, x_2, \dots, x_M) = C \times \pi_1(x_1) \times \pi_2(x_2) \times \dots \times \pi_M(x_M)$  (**απόδειξη με επαλήθευση εξισώσεων ισορροπίας μεταβάσεων**)