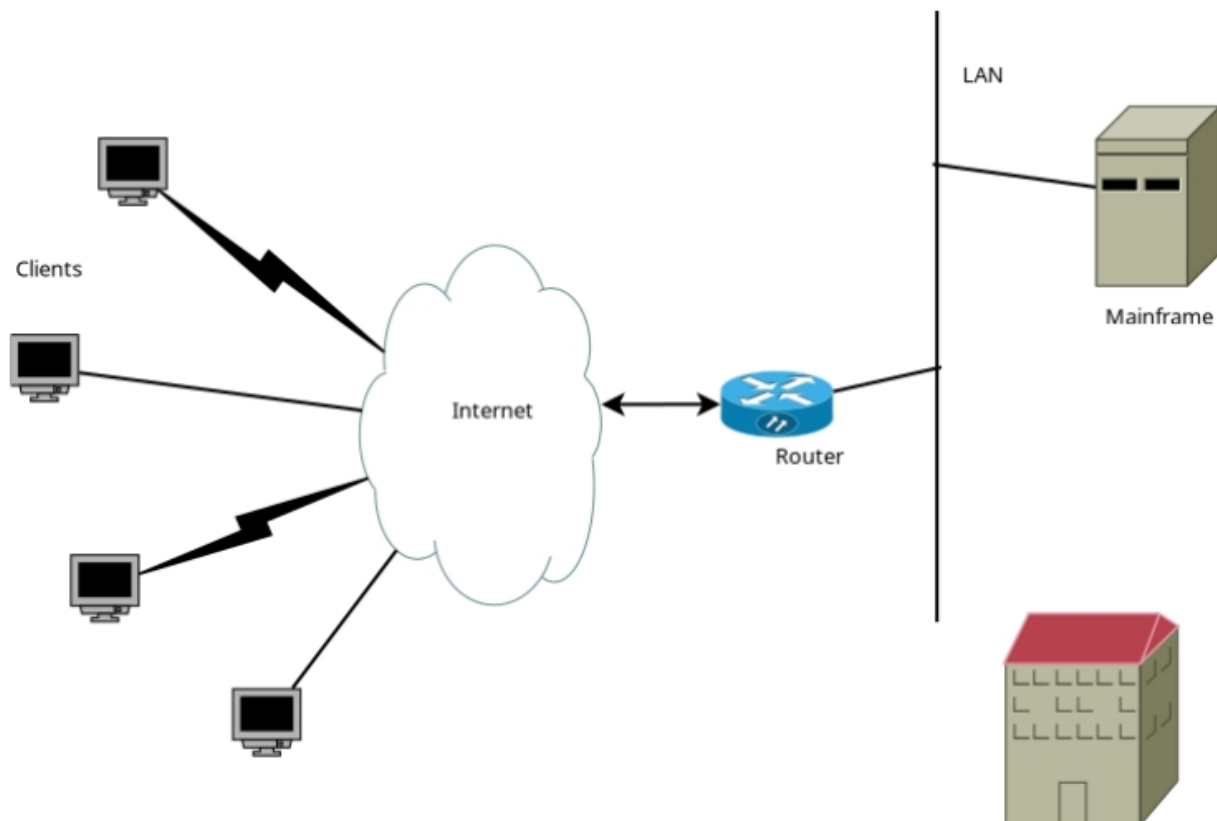


Εθνικό Μετσόβειο Πολυτεχνείο
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών
Επίδοση Υπολογιστικών Συστημάτων – 8ο Εξάμηνο
3η Άσκηση – Ακ. Έτος 2021-2022

Ομάδα 66
Κυριακόπουλος Γεώργιος – el18153
Τζελέπης Σεραφείμ – el18849

Εισαγωγή

Σε αυτή την εργασία θα ασχοληθούμε με την επίδοση ενός υπολογιστικού συστήματος ενός οργανισμού παροχής ιατρικών υπηρεσιών. Συγκεκριμένα, το τμήμα ενημέρωσης αυτού του οργανισμού υποστηρίζεται από έναν κεντρικό εξυπηρετητή Mainframe, ο οποίος αποτελείται από μία ΚΜΕ (CPU) και δύο δίσκους. Η γενικότερη διασύνδεση του συστήματος καθώς και ο τρόπος επικοινωνίας των πελατών με το σύστημα φαίνεται στο παρακάτω σχήμα.

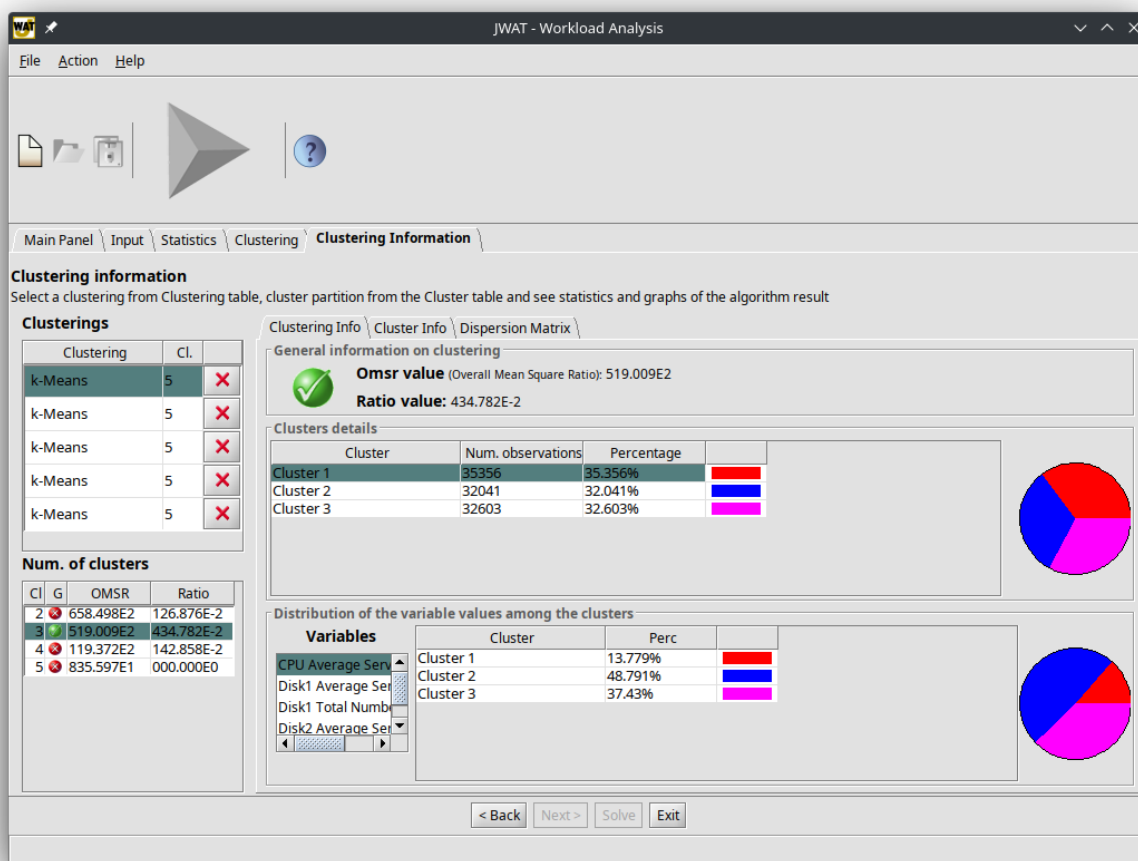


A. Χαρακτηρισμός φορτίου

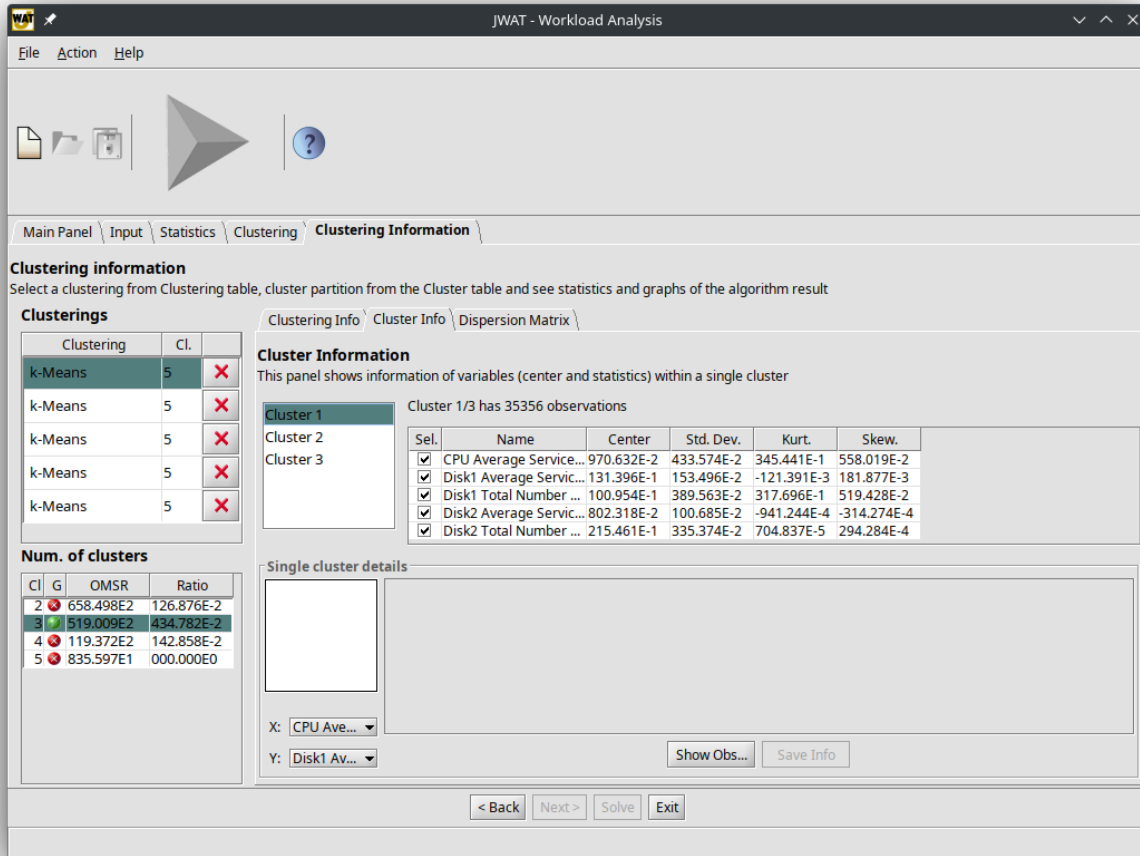
Αρχικά, θα πρέπει να προσδιορίσουμε τις παραμέτρους του φορτίου, με βάση το δοσμένο αρχείο *server.log*. Θα χρησιμοποιήσουμε τον αλγόριθμο k-means για το clustering, έχοντας λάβει νωρίτερα τυχαία δειγματοληψία 100.000 μετρήσεων (σε σύνολο 298.090 μετρήσεων). Ως παραμέτρους ομαδοποίησης για το clustering έχουμε:

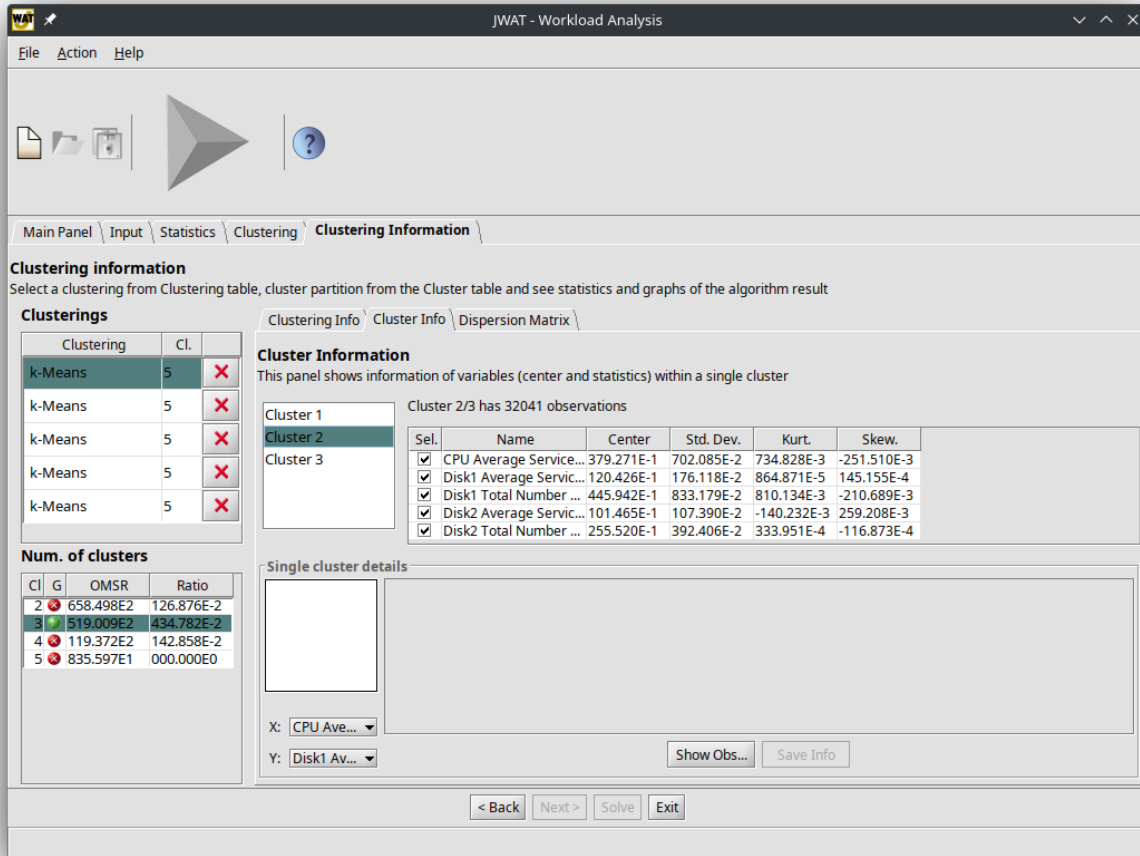
- Μέγιστος αριθμός ομάδων: $k=5$
- Αριθμός επαναλήψεων (iterations): 50
- Μετασχηματισμός (transformation): $(\text{value} - \text{min}) / (\text{max} - \text{min})$

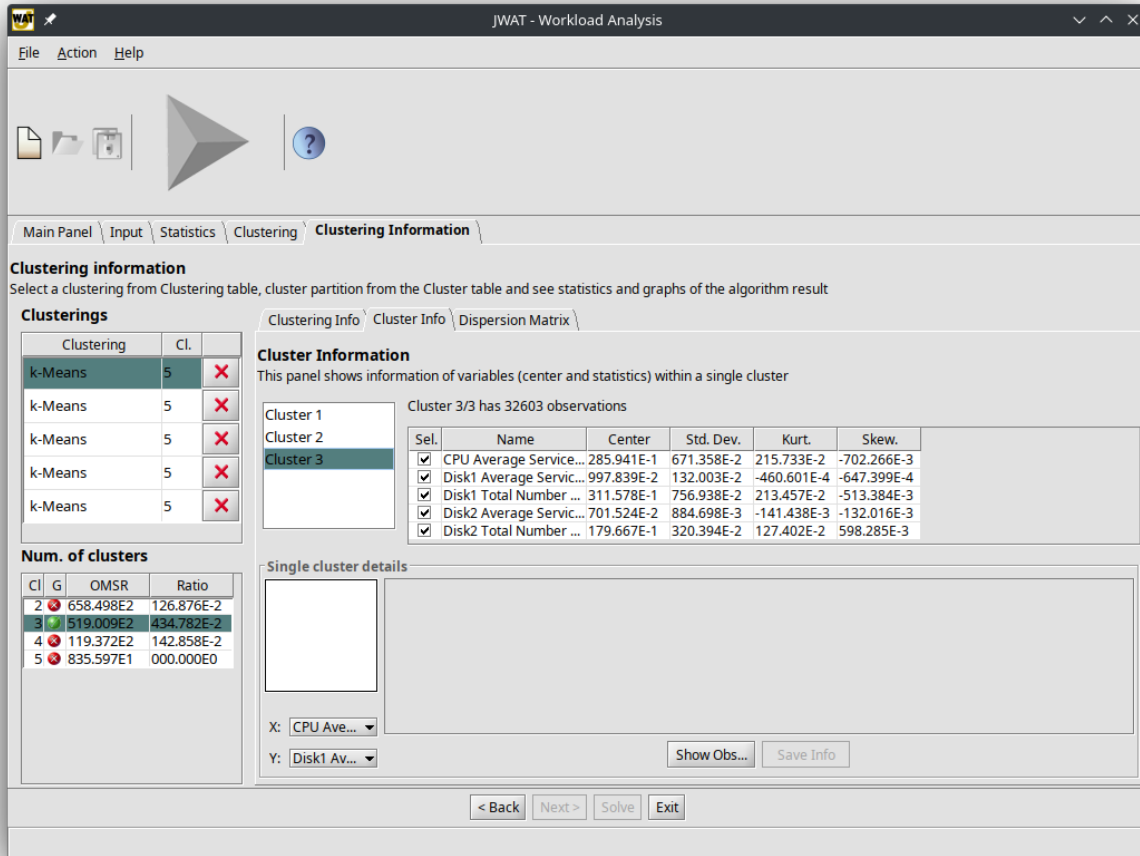
Τρέξαμε τον αλγόριθμο 5 φορές, όπως ζητήθηκε και τελικά επιλέξαμε το clustering με αριθμό ομάδων $k=3$ και με το μεγαλύτερο Ratio. Το επιλεγμένο clustering φαίνεται στην παρακάτω εικόνα, με τα ποσοστά για κάθε κατηγορία κάτω από το Clusters Details.



Ακολουθούν 3 εικόνες με τα Cluster Info για τις 3 κατηγορίες εργασιών Cluster 1, Cluster 2, Cluster 3, όπου φαίνονται ο μέσος χρόνος εξυπηρέτησης για το CPU, Disk1, Disk2 και ο συνολικός αριθμός επισκέψεων στους σταθμούς Disk1, Disk2 ανά εργασία.







B. Μοντέλο ανοιχτού δικτύου

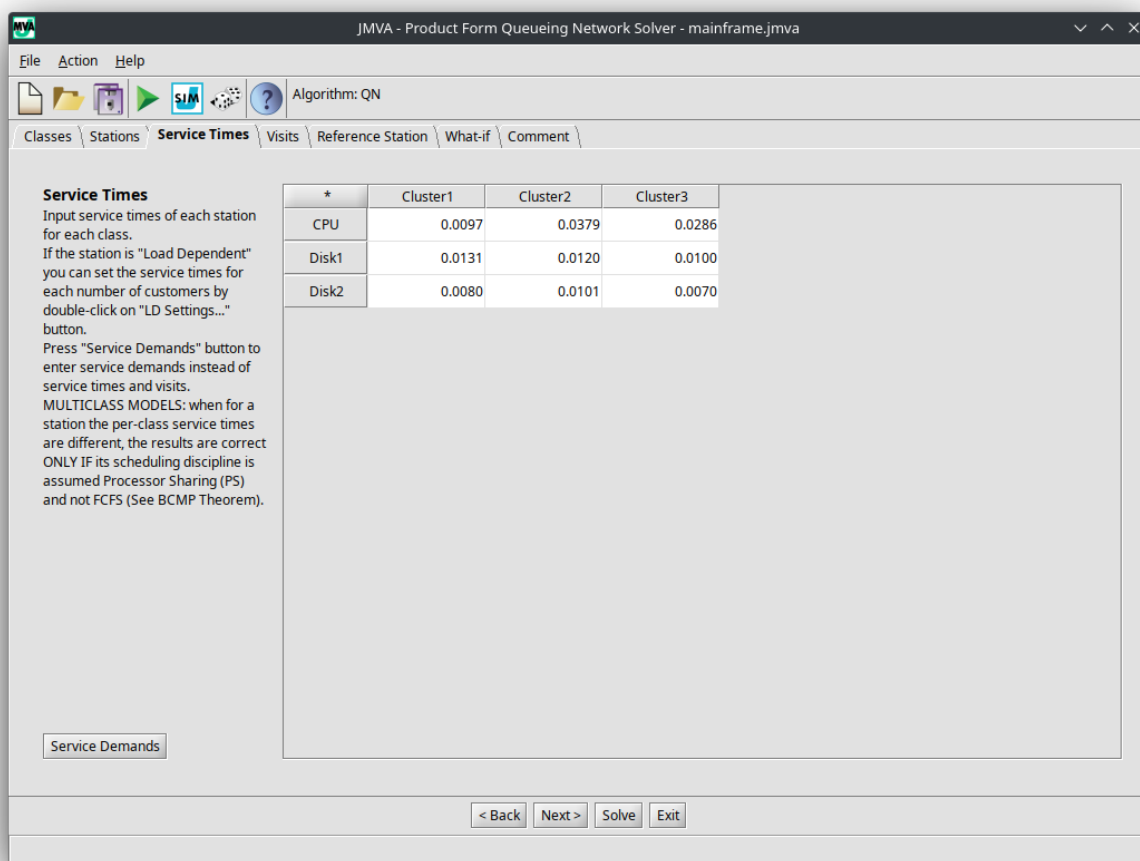
Με βάση το Clusters Details, θα προσδιορίσουμε τους διάφορους ρυθμούς αφίξεων. Αρχικά όμως πρέπει να προσδιοριστεί ο συνολικός ρυθμός αφίξεων του δικτύου. Το συνολικό λ θα είναι

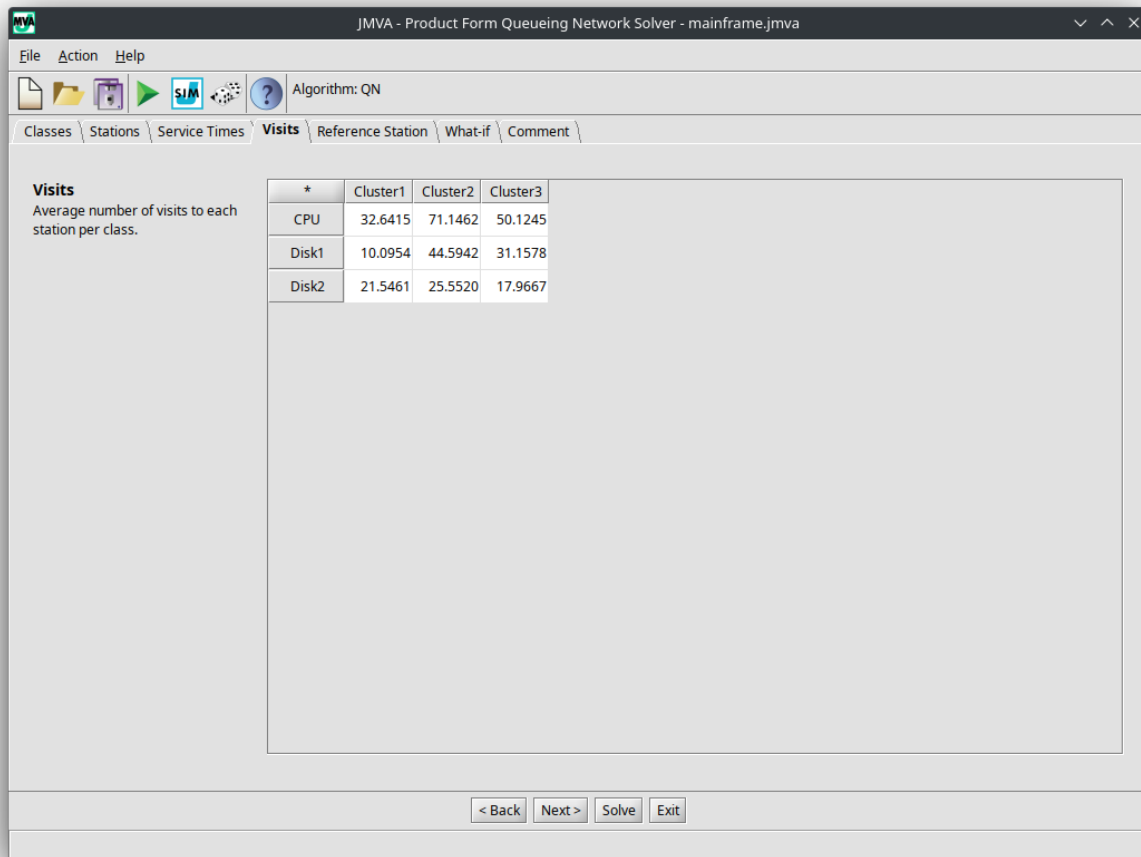
$$\lambda = \frac{298090}{512000} = 0.5822 \text{ } \epsilon \rho \gamma / \text{sec} .$$

Άρα, ο ρυθμός αφίξεων της κάθε κατηγορίας θα είναι ίσος με το συνολικό ρυθμό αφίξεων επί του ποσοστού της κατηγορίας. Συγκεκριμένα:

- $\lambda_1 = 0.5822 \cdot 0.35356 = 0.2058 \text{ } \epsilon \rho \gamma / \text{sec} ,$
- $\lambda_2 = 0.5822 \cdot 0.32041 = 0.1865 \text{ } \epsilon \rho \gamma / \text{sec} ,$
- $\lambda_3 = 0.5822 \cdot 0.32603 = 0.1898 \text{ } \epsilon \rho \gamma / \text{sec} .$

Στην συνέχεια θα καταγράψουμε το μέσο χρόνο εξυπηρέτησης ανά επίσκεψη και το μέσο αριθμό επισκέψεων για κάθε σταθμό και κατηγορία, με βάση τα δεδομένα και τα χαρακτηριστικά των τριών κατηγοριών (Cluster1, Cluster2, Cluster3).



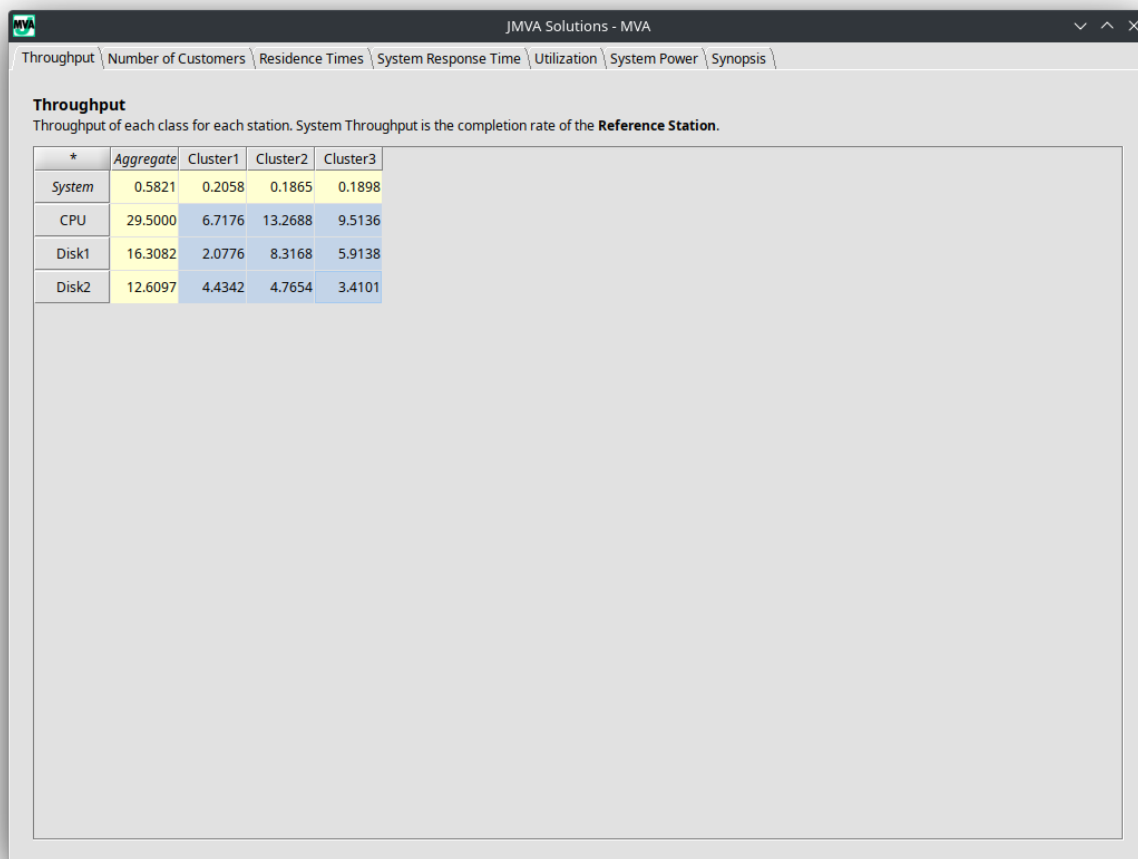


Τα δεδομένα για τους μέσους χρόνους εξυπηρέτησης των 3 σταθμών, όπως και για τους μέσους αριθμούς επισκέψεων για τους δίσκους 1, 2 λαμβάνονται από τα αποτελέσματα του Cluster Info για κάθε cluster, ενώ ο μέσος αριθμός επισκέψεων για την ΚΜΕ (CPU) προκύπτει ως εξής

$$v_{cpu} = v_{disk1} + v_{disk2} + 1.$$

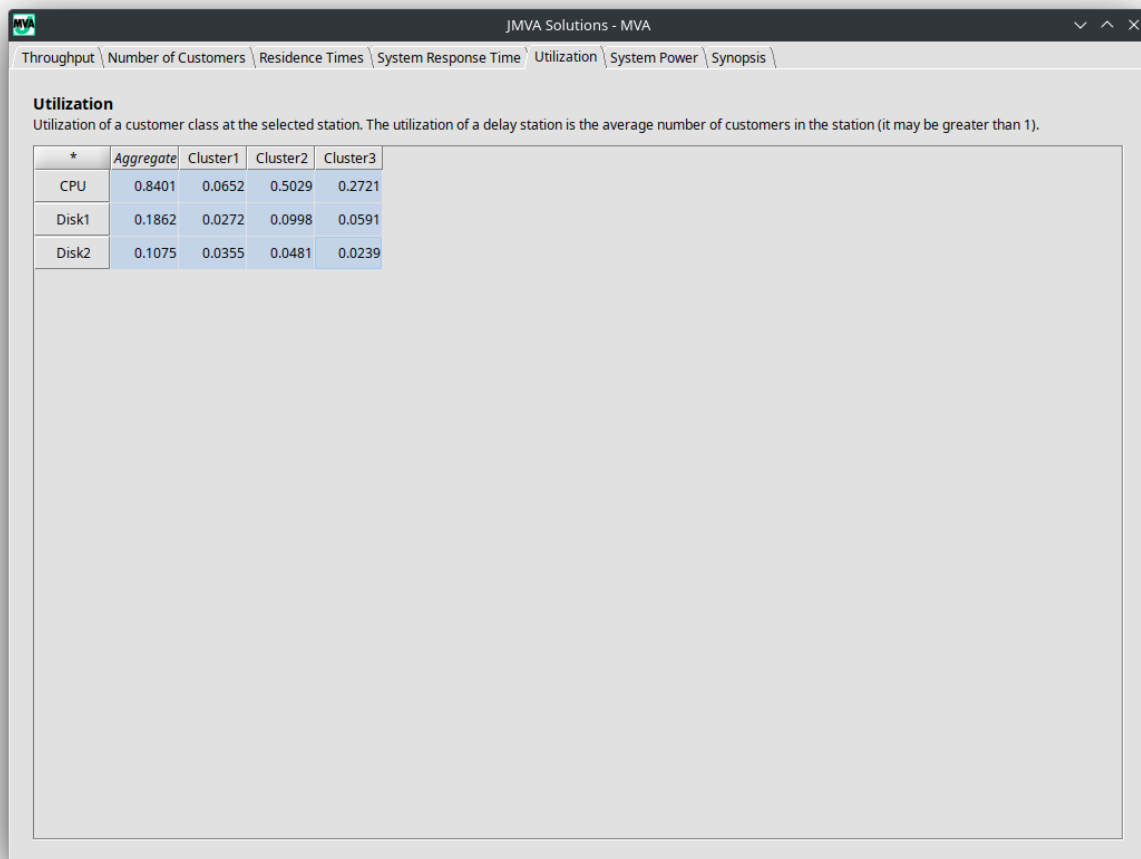
Με τη χρήση του εργαλείου JMVA του JMT, έχουμε για κάθε σταθμό και ανά κατηγορία (οι επιλεγμένες τιμές – με μπλε χρώμα):

- Ρυθμός απόδοσης



*	Aggregate	Cluster1	Cluster2	Cluster3
System	0.5821	0.2058	0.1865	0.1898
CPU	29.5000	6.7176	13.2688	9.5136
Disk1	16.3082	2.0776	8.3168	5.9138
Disk2	12.6097	4.4342	4.7654	3.4101

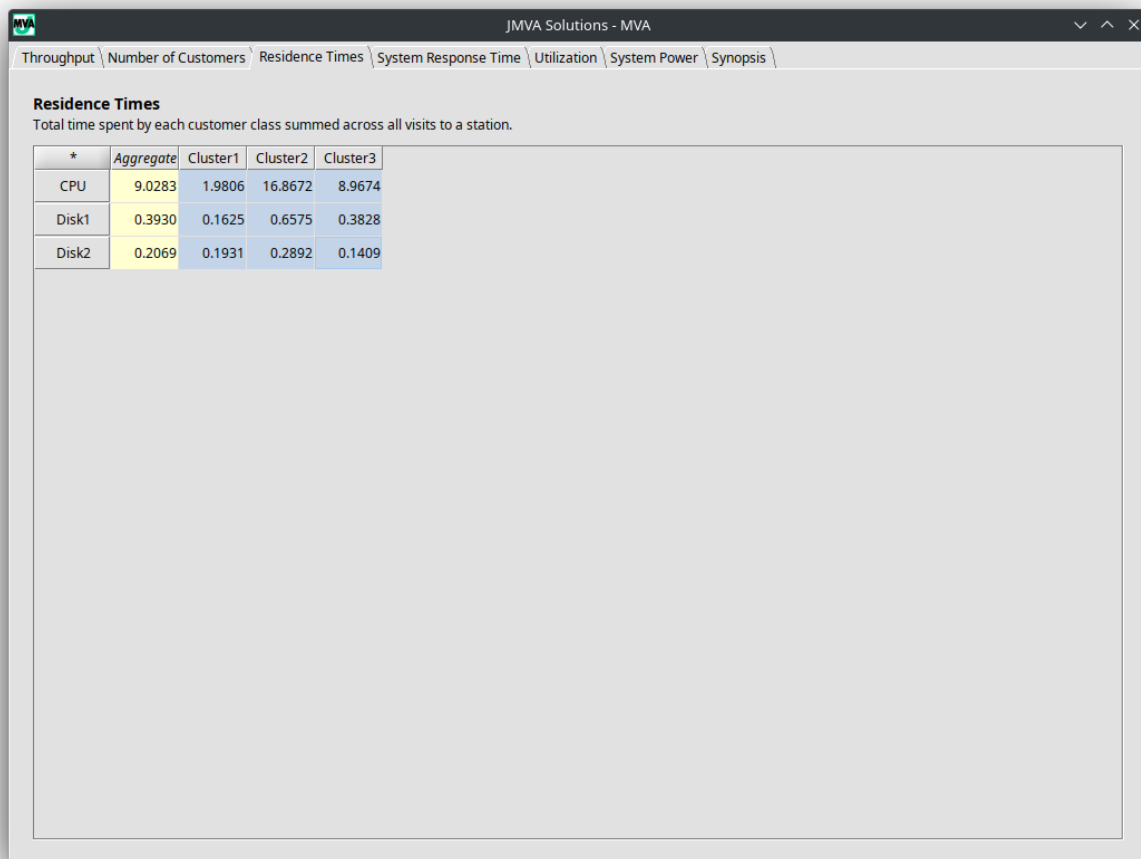
- Βαθμός χρησιμοποίησης (ανά κατηγορία και συνολικά για κάθε σταθμό)



The screenshot shows a software window titled "JMVA Solutions - MVA". It has a menu bar with the following items: Throughput, Number of Customers, Residence Times, System Response Time, Utilization (which is highlighted), System Power, and Synopsis. Below the menu bar, the "Utilization" section is active. It contains a title "Utilization" and a descriptive text: "Utilization of a customer class at the selected station. The utilization of a delay station is the average number of customers in the station (it may be greater than 1).". Below this text is a table with utilization data.

*	Aggregate	Cluster1	Cluster2	Cluster3
CPU	0.8401	0.0652	0.5029	0.2721
Disk1	0.1862	0.0272	0.0998	0.0591
Disk2	0.1075	0.0355	0.0481	0.0239

- Μέσος χρόνος απόκρισης



The screenshot shows a software window titled "JMVA Solutions - MVA" with a menu bar containing "Throughput", "Number of Customers", "Residence Times", "System Response Time", "Utilization", "System Power", and "Synopsis". The "Residence Times" menu item is selected. Below the menu bar, the text "Residence Times" is displayed, followed by a description: "Total time spent by each customer class summed across all visits to a station." A table with 5 columns is shown: "*", "Aggregate", "Cluster1", "Cluster2", and "Cluster3". The table contains three rows of data: "CPU", "Disk1", and "Disk2". The "Aggregate" column for each row is highlighted in yellow, and the "Cluster" columns are highlighted in blue.

*	Aggregate	Cluster1	Cluster2	Cluster3
CPU	9.0283	1.9806	16.8672	8.9674
Disk1	0.3930	0.1625	0.6575	0.3828
Disk2	0.2069	0.1931	0.2892	0.1409

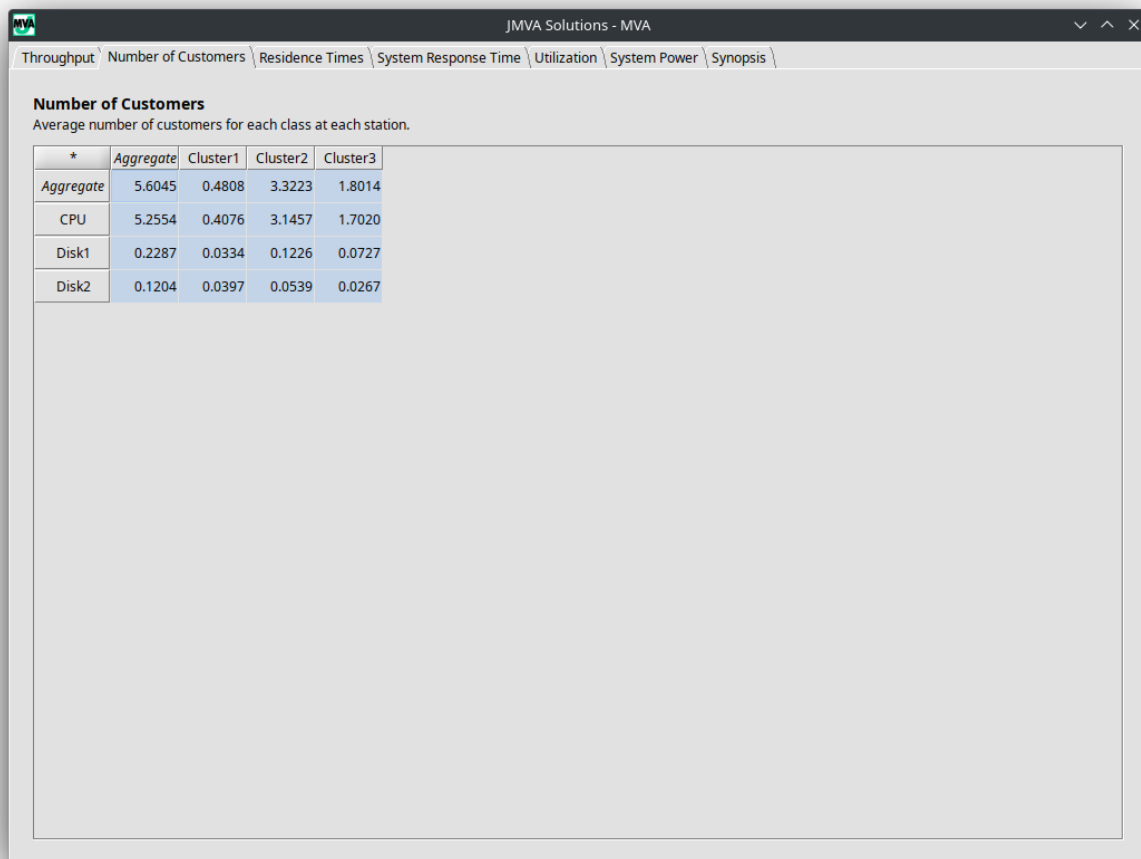
Στη συνέχεια, έχουμε συνολικά για το δίκτυο και ανά κατηγορία:

- Μέσος χρόνος απόκρισης

System Response Time
The global aggregate is the "System Response Time" and is obtained weighting the aggregated values by the relative per-class throughput.
A: This value of System Response Time **includes** the Residence Time of the Reference Station.
B: This value of System Response Time **does NOT include** the Residence Time of the Reference Station.
Notice: For **open classes** the Reference Station always coincides with the arrival process. Thus the **B** values are not computed.

*	Aggregate	Cluster1	Cluster2	Cluster3
A	9.6281	2.3362	17.8139	9.4912
B	--	--	--	--

- Μέσος αριθμός εργασιών



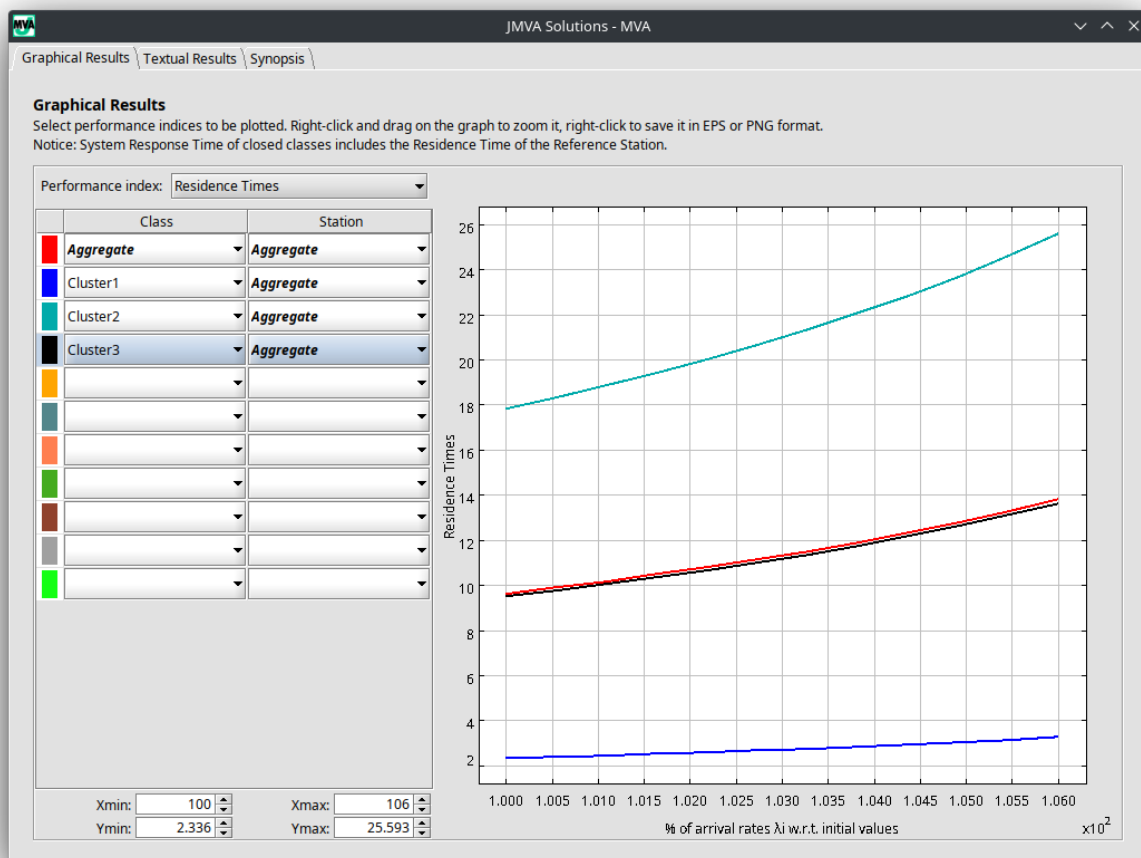
Number of Customers
Average number of customers for each class at each station.

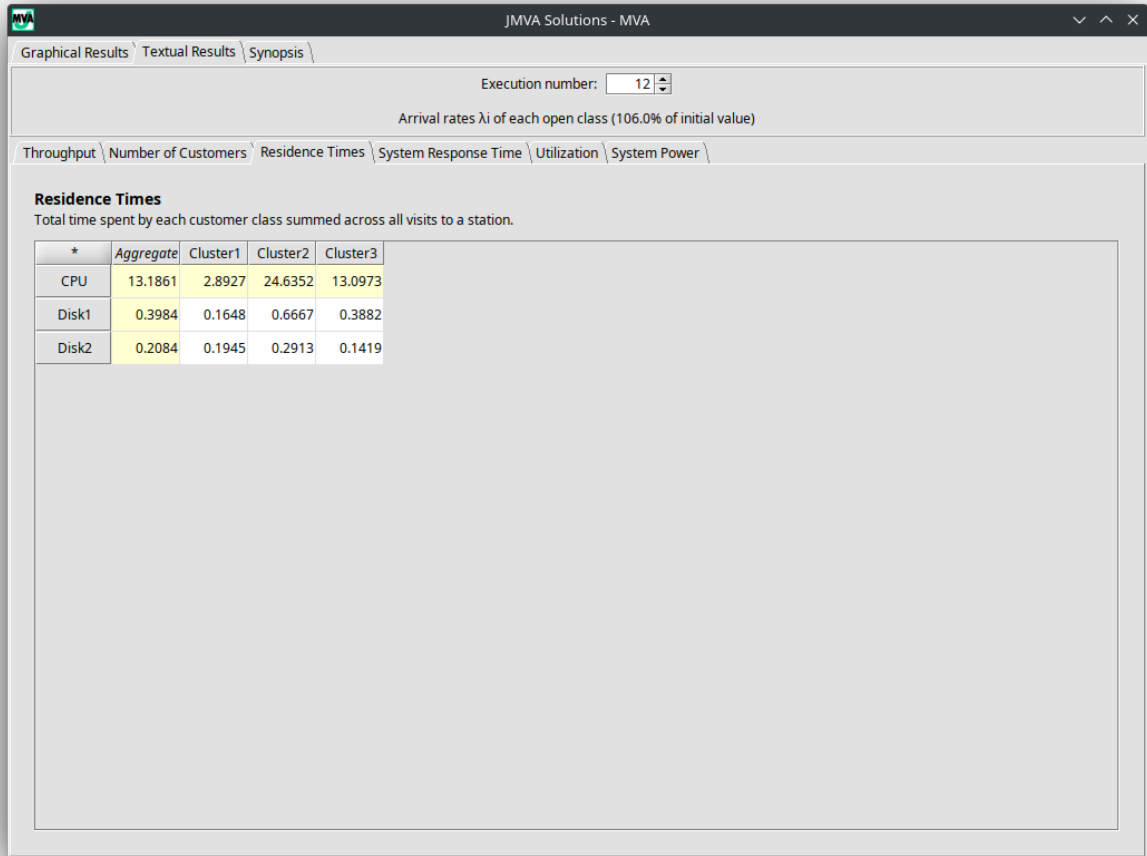
*	Aggregate	Cluster1	Cluster2	Cluster3
Aggregate	5.6045	0.4808	3.3223	1.8014
CPU	5.2554	0.4076	3.1457	1.7020
Disk1	0.2287	0.0334	0.1226	0.0727
Disk2	0.1204	0.0397	0.0539	0.0267

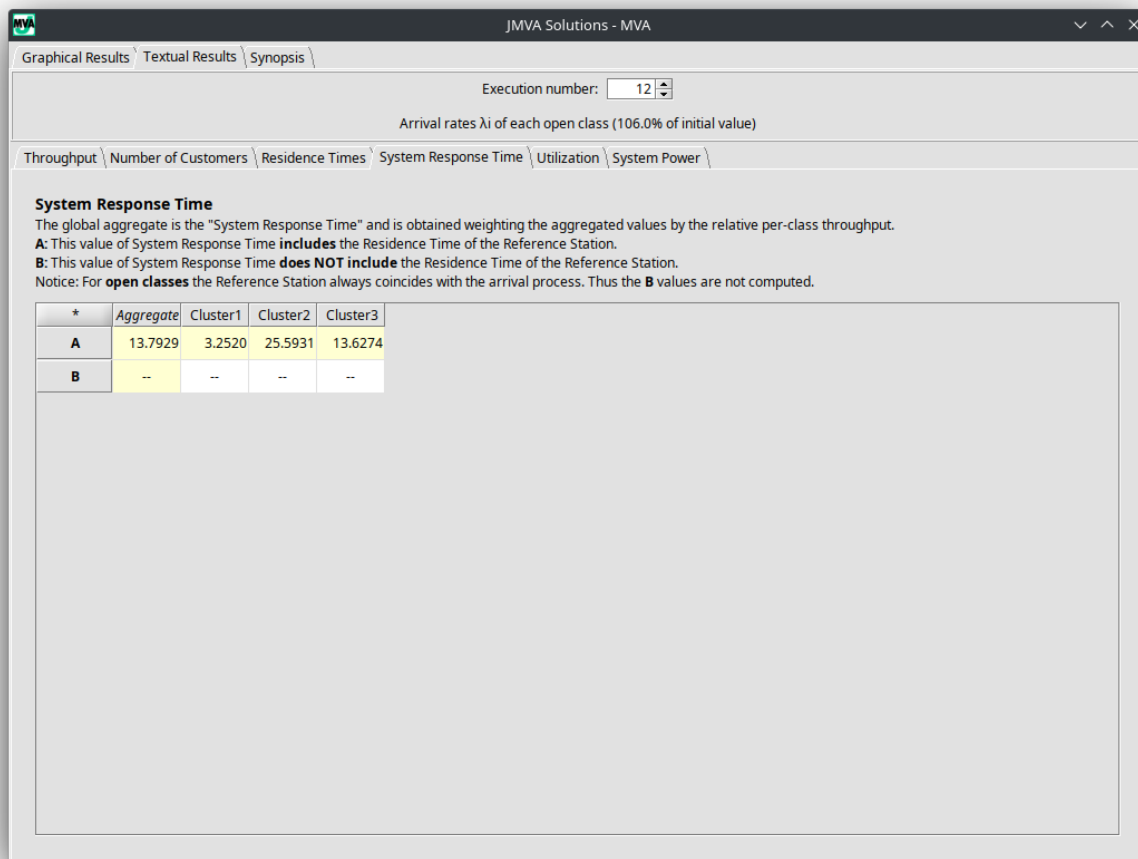
Γ. Πρόβλεψη επίδοσης

Θέλουμε να μελετήσουμε το σύστημα για τους επόμενους 12 μήνες λειτουργίας, όπου θα υπάρχει αύξηση των ρυθμών αφίξεων, αναλογικά για τις τρεις κατηγορίες, κατά +0.5% κάθε μήνα, σε σχέση με την αρχική κατάσταση του συστήματος. Δηλαδή, θα υπάρξει μία συνολική αύξηση 6% στο τέλος του διαστήματος του ενός έτους.

Χρησιμοποιώντας το what-if analysis εργαλείο του JMT, λαμβάνουμε αρχικά τους χρόνους απόκρισης των αιτημάτων για όλες τις κατηγορίες (μελετάμε για $n=12$, δηλαδή, ρυθμούς αφίξεων αυξημένους κατά 6%).







Παρατηρούμε ότι για τα Cluster1, Cluster3 ο χρόνος απόκρισης των αιτημάτων είναι εντός του αποδεκτού ορίου (μικρότερο των 15 sec), ενώ για το Cluster2 ο αντίστοιχος χρόνος είναι εκτός ορίων (25.5931 sec). Επομένως, χρειάζεται να γίνει κάποια αλλαγή στα συστατικά του Mainframe.

Βλέποντας τους επιμέρους χρόνος απόκρισης σε κάθε σταθμό για το Cluster2, παρατηρούμε ότι το πρόβλημα είναι η ΚΜΕ (CPU), καθώς οι δίσκοι έχουν χρόνους που και να μηδενιστούν δεν θα πέσει ο συνολικός χρόνος απόκρισης κάτω από τα 15 sec. Επομένως, θα χρειαστεί να αναβαθμίσουμε την ταχύτητα της ΚΜΕ. Συμβολίζουμε με 1 το σταθμό ΚΜΕ και με 1, 2, 3 τις κατηγορίες Cluster1, Cluster2, Cluster3, αντίστοιχα. Επίσης, με $'$ συμβολίζουμε τις νέες τιμές που αντιστοιχούν στη νέα, ταχύτερη ΚΜΕ.

Θεωρώντας ότι οι χρόνοι απόκρισης των δίσκων θα παραμείνουν σταθεροί, για τους χρόνους απόκρισης της ΚΜΕ θα ισχύουν

$$\frac{S'_{1j}}{S_{1j}} = x \quad (1),$$

$$U'_{ij} = \lambda_j \cdot D'_{ij} = \lambda_j \cdot S'_{ij} \cdot v_{ij} = \lambda_j \cdot S_{ij} \cdot x \cdot v_{ij} = U_{ij} \cdot x \quad (2)$$

και επομένως για τις ταχύτητες της ΚΜΕ

$$\frac{CPU'_{speed}}{CPU_{speed}} = \frac{1}{x} \quad (3)$$

και έστω οριακή τιμή απόκρισης τα 15 sec, βρίσκουμε αρχικά τη μέγιστη τιμή χρόνου απόκρισης για τη ΚΜΕ.

$$R^2=15 \rightarrow R_{11}+R_{12}+R_{13}=15 \rightarrow R_{12}+0.6667+0.2913=15 \rightarrow R_{12}=14.042 \quad (4)$$

$$R_{12}=\frac{D'_{12}}{1-\sum_{k=1}^3 U'_{1k}} \rightarrow R_{12}=\frac{S'_{12} \cdot v_{12}}{1-(U'_{11}+U'_{12}+U'_{13})} \quad (1),(2) \rightarrow$$

$$R_{12}=\frac{S_{12} \cdot x \cdot v_{12}}{1-(U_{11} \cdot x+U_{12} \cdot x+U_{13} \cdot x)} \rightarrow R_{12}=\frac{S_{12} \cdot x \cdot v_{12}}{1-U_1 \cdot x} \quad (5)$$

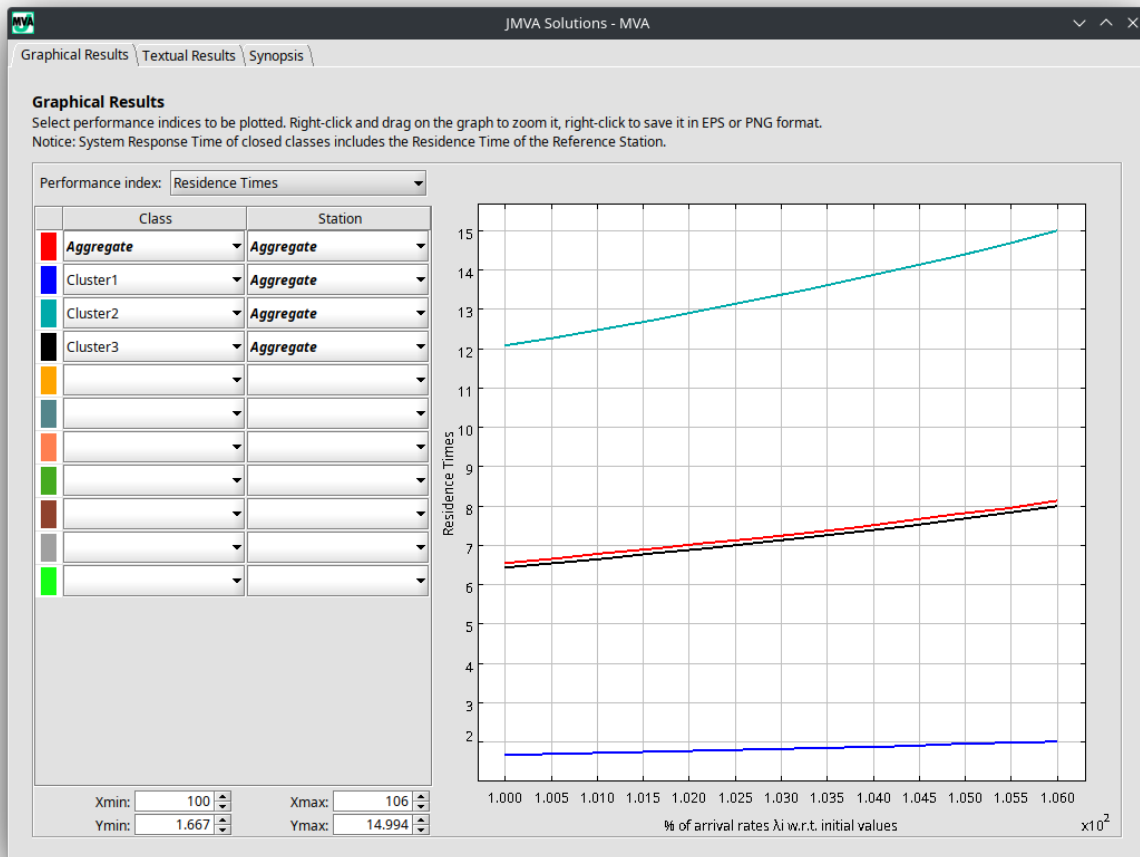
Επομένως, από τις σχέσεις (4) και (5) παίρνουμε:

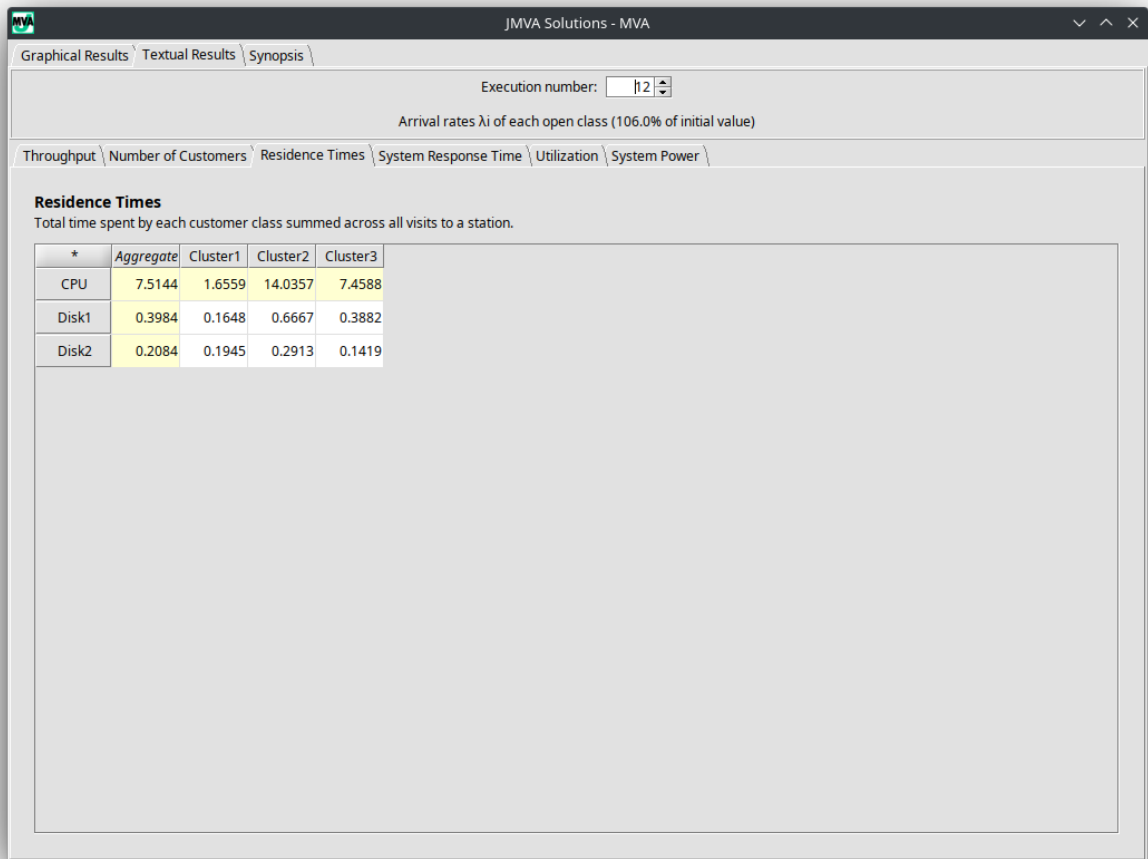
$$\begin{aligned} \frac{S_{12} \cdot x \cdot v_{12}}{1-U_1 \cdot x} &= 14.042 \rightarrow S_{12} \cdot x \cdot v_{12} = 14.042 - 14.042 \cdot U_1 \cdot x \rightarrow x = \frac{14.042}{S_{12} \cdot v_{12} + 14.042 \cdot U_1} \\ \rightarrow x &= \frac{14.042}{0.0379 \cdot 71.1462 + 14.042 \cdot 0.8905} \rightarrow x = \frac{14.042}{15.2} \rightarrow x = 0.924 \rightarrow \\ \frac{1}{x} &= 1.082 \end{aligned}$$

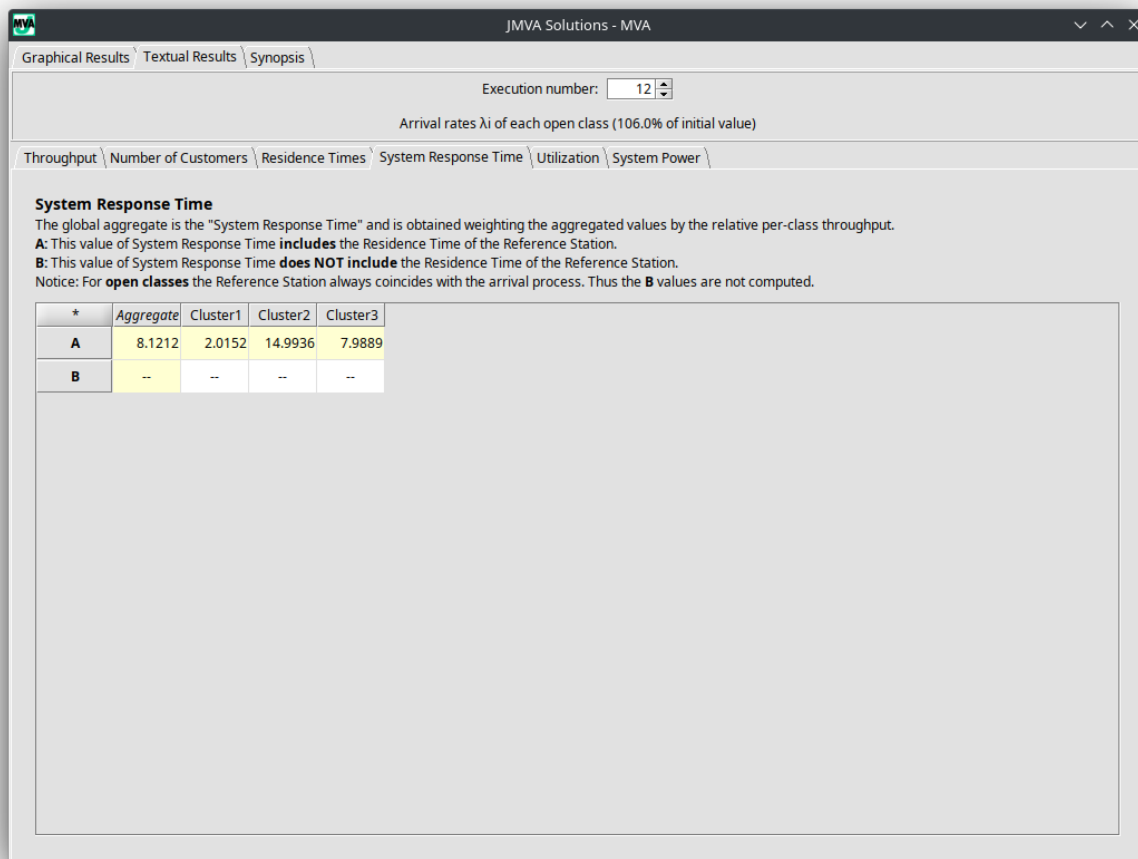
Επομένως, με μία αύξηση της ταχύτητας της ΚΜΕ κατά 8.2% θα έχουμε το επιθυμητό χρόνο απόκρισης και στο Cluster2, όπως φαίνεται και με την νέα προσομοίωση παρακάτω, με τους νέους μέσους χρόνος εξυπηρέτησης

$$S'_{1j} = S_{1j} \cdot x = S_{ij} \cdot 0.924,$$

δηλαδή $S'_{11}=0.009, S'_{12}=0.035, S'_{13}=0.0264$.







Βλέπουμε, επομένως, ότι με αυτήν την αλλαγή στην ταχύτητα της ΚΜΕ (CPU), επιτυγχάνεται το ζητούμενο, δηλαδή, οι χρόνοι απόκρισης του συστήματος για τα αιτήματα όλων των κατηγοριών να είναι μικρότεροι των 15 sec.