

Εθνικό Μετσόβιο Πολυτέχνειο

Επίδοση Υπολογιστικών Συστημάτων

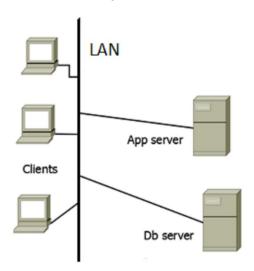
1η Εργασία

Αχλάτης Στέφανος-Σταμάτης (03116149) < el16149@central.ntua.gr>

1.Εισαγωγή

Στο παρόν Project μελετάται ένα σύστημα εξυπηρετητή-πελατών ως ανοιχτό δίκτυο. Πιο συγκεκριμένα μελετάμε το ιδιωτικό δίκτυο LAN ενός ειδησεογραφικού οργανισμού. Στην μελέτη μας θεωρούμε ότι ο οργανισμός έχει 480 χρήστες αυτού του δικτύου, και οι δυο λειτουργίες που μπορούν να εκτελέσουν είναι εγγραφές και ανανεώσεις.

Το σύστημα ακολουθεί αρχιτεκτονική δύο βαθμίδων η οποία περιλαμβάνει εξυπηρετητή εφαρμογών (App Server) και εξυπηρετητή βάσης δεδομένων (Db Server). Οι εξυπηρετητές περιλαμβάνουν CPU και έναν δίσκο έκαστος.



Και μας έχουν δώσει τον εξής αρχικό πίνακα παραμέτρων του συστήματος:

			D _{ij} (msec)
α/α	Σταθμός	Τύπος	Εγγραφές	Ανανεώσεις
1	Term	D	25000	19000
2	LAN	D	47	24
3	App Server CPU	LI	43	19
4	App Server Disk	LI	44	29
5	Db Server CPU	LI	45	59
6	Db Server Disk	LI	41	37

Αξίζει να σημειωθεί ότι ο σταθμός term δηλώνει τον χρόνο που δαπανά ο χρήστης του συστήματος για να σκεφτει και να αρχικοποίηση τις απαιτούμενες πληροφορίες στο σύστημα έτσι ώστε το σύστημα να εκτελέσει την απαιτούμενη ενέργεια. Τονίζεται ότι αυτός ο χρόνος δεν περιλαμβάνει τον χρόνο απόκρισης του συστήματος, δηλαδή πόσο χρόνο κάνει το σύστημα να αναζητήσει, τροποποιήσει κλπ και να δώσει στον χρήστη αυτές τις πληροφορίες.

Γι αυτόν τον λόγο αυτός ο χρόνος είναι της τάξεως των αρκετων δευτερολέπτων σε αντίθεση με τους υπόλοιπους χρόνους που είναι της τάξεως των msec.

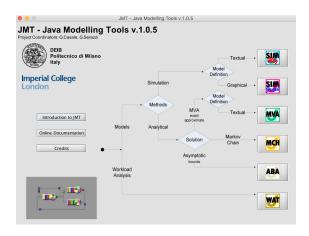
Έτσι, στην αρχή ψάχνουμε να βρούμε το μειγμα πληθυσμού που μεγιστοποιει το throughput του συστηματος ως σύνολο μέσω του εργαλείου what-if του JMVA, και για το δεδομένο μειγμα μελετάμε ορισμένες μετρικές και καταλήγουμε στο συμπέρασμα ότι οι αρχικές προδιαγραφές της cpu του db server δημιουργεί στένωση και καθυστέρηση στο σύστημα. Γι

αυτό στην συνεχεια αντικαθιστούσε την cpu του db server με μια load dependent cpu και βλέπουμε πως επηρεάζει το σύστημα.

Στην συνέχεια, με στόχο την απλούστευση της προσομοίωσης, ομαδοποιούμε το προηγούμενο σύστημα σε μοντέλο μιας κατηγορίας όπου η τιμή της απαίτησης εξυπηρέτησης σε κάθε σταθμό του μοντέλου μιας κατηγορίας θα τεθεί ίση με τον μέσο όρο των επιμέρους τιμών σταθμισμένων με τον ρυθμό απόδοσης της αντίστοιχης κατηγορίας. Τέλος, γι αυτό το νέο απλοποιημένο σύστημα παρατηρούμε ότι η cpu του db server δημιουργεί στένωση και γι αυτό την αντικαθιστούμε με μια load dependent cpu και βλέπουμε πως αυτό επηρεάζει το σύστημα.

2.Το εργαλείο JMVA

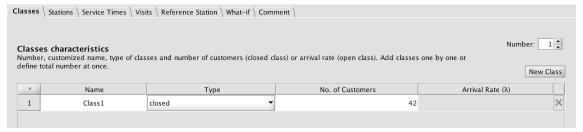
Οι προσομοιώσεις έγιναν στο εργαλείο JMVA του πακέτου JMT.



Το εργαλείο αυτό εφαρμόζει τον Mean Value Algorithm σε ένα σύστημα το οποίο ορίζει ο χρήστης.

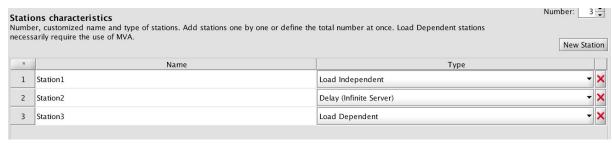
Αρχικά ο χρήστης μπορεί να ορίσει το πληθυσμό της προσομοίωσης, χωρίζοντας το σε κλάσεις αν χρειάζεται ορίζοντας την καθε μια ως κλειστή ή ανοικτή.

Για το συγκεκριμένο πρόβλημα οι κλάσεις μας ήταν κλειστές αφού δεν επιτρέπεται νεα ροη στο σύστημα πέρα από του 480 χρήστες του.

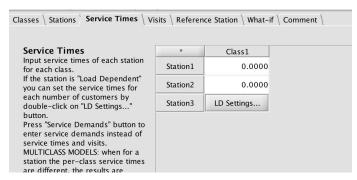


Μετά στην συνέχεια ο χρήστης μπορεί να ορίσει τους σταθμούς του συστήματος, και να τους ονοματοδοτήσει έτσι ώστε να μπορεί να τους ξεχωρίσει ευκολότερα.

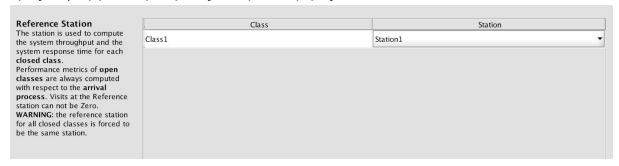
Επίσης μπορεί να τους χαρακτηρίσει ως load dependent, load independent και ως delay station.



Στη συνέχεια μπορεί να ορίσει τους χρόνους κάθε σταθμού και για την περίπτωση των load dependent station να δώσει καθυστέρηση ανάλογα με το πλήθος χρηστών στο συγκεκριμένο station.



Μετά μπορεί να ορίσει σταθμό αναφοράς ,με βάση τον οποίον το jmva υπολογίζει την ρυθμαπόδοση και το system response time για κάθε κλειστή κλάση. Εμείς στην εργασία ορίσαμε ως σταθμό αναφοράς τον term.



Τέλος ο χρήστης μπορεί να τρέξει what-if analysis ως προς πολλές μετρικές. Εμείς τρέξαμε ως προς το πλήθος των χρηστών έτσι ώστε να βρούμε το μείγμα πληθυσμού που βελτιστοποιεί την ρυθμαπόδοση του συστηματος.



Τέλος ο χρήστης μπορεί να επιλέξει να τρέξει διάφορες παραλλαγές του mva αλλά εμείς επιλέξαμε να τρέξουμε τον default mva όπως αυτός διδάχτηκε στο μάθημα.

Κλείνοντας, αξίζει να τονιστεί ότι αντίστοιχα την προσομοίωση θα μπορουσαμε να την τρέχαμε γράφοντας κώδικα σε γλώσσες όπως η python3, η C++ ή ακόμη και η Java όπου

θα κάναμε πολύ συγκεκριμένη κάθε τεχνική λεπτομέρεια του συστήματος χωρίς να αναγκαζόμαστε να προσαρμόσουμε το σύστημα στις ανάγκες του προσομοιωτή. Αλλά όπως μάθαμε και στο μάθημα αποτελεί καλή τεχνική πρώτα να τρέχουμε τις προσομοιώσεις μας σε ενα generic περιβάλλον όπως ο JMVA και μετά να χτίζουμε εμείς οι ίδιοι δικές μας προσομοιώσεις για να μελετήσουμε με ακόμη μεγαλύτερη ακρίβεια το σύστημα μας.

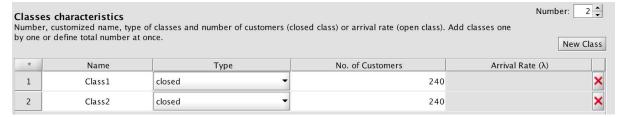
Έτσι, σε αυτή την εργασία θα τρέξουμε τις προσομοιώσεις μόνο μέσα από το JMVA.

3.Εκτέλεση Άσκησης Ερώτημα Α

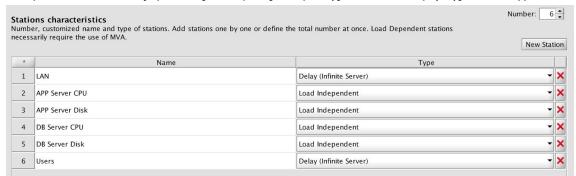
Με χρήση της επιλογής What-if του εργαλείου να προσδιοριστεί το μείγμα πληθυσμού των δύο κατηγοριών, για το οποίο επιτυγχάνεται μεγιστοποίηση του ρυθμού απόδοσης του συστήματος. Το μείγμα αυτό θα ισχύει στα επόμενα ερωτήματα.

Ζητούνται ο ρυθμός απόδοσης και ο χρόνος απόκρισης του συστήματος συνολικά και για καθεμία από τις κατηγορίες πελατών, καθώς και ο βαθμός χρησιμοποίησης των σταθμών του συστήματος. Ποιος σταθμός είναι η στένωση του συστήματος;

Αρχικά, δηλώνουμε τις 2 κατηγορίες χρηστών με τους αντίστοιχους πληθησιμούς. Η κατηγορία 1 ανήκει στις εγγραφές ενώ η κατηγορία 2 ανήκει στις ανανεώσεις. Δεδομένου ότι ο στόχος της προσομοίωσης είναι να μελετήσουμε το population mix βάζουμε δύο αριθμούς στον πληθυσμό της κάθε κατηγορίας, τέτοιοι ώστε **το άθροισμα** τους να είναι ίσο με το μέγιστο δυνατό πληθυσμό που ορίζεται στο σύστημα ως σύνολό, δηλαδή 480. Επιλέγουμε να αρχικοποιήσουμε τον πληθυσμό σε 240 σε κάθε κατηγορία.



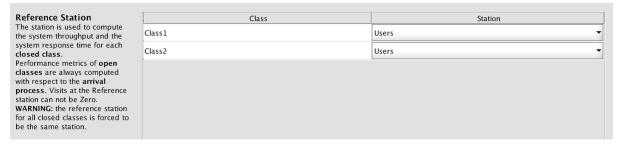
Μετά σειρά έχει η δήλωση των σταθμών του συστήματος και ο χαρακτηρισμός τους ως Load Independent και Delay για τους σταθμούς αναμονής και καθυστέρησης αντίστοιχα.



Στη συνέχεια ορίσαμε τις απαιτήσεις σε μέσο χρόνο χρήστη,δηλαδή σε μέσο χρόνο σκέψης, και σε μέση συνολική απαίτηση εξυπηρέτησης για κάθε σταθμό. Οι τιμές που φαίνονται στον πίνακα είναι της τάξεως των sec.

Service Times	*	Class1	Class2
Input service times of each station for each class.	LAN	0.0470	0.0240
If the station is "Load Dependent" you can set the service times for	APP Server	0.0430	0.0190
each number of customers by double-click on "LD Settings"	APP Server	0.0440	0.0290
button. Press "Service Demands" button to	DB Server	0.0450	0.0590
enter service demands instead of service times and visits.	DB Server	0.0410	0.0370
MULTICLASS MODELS: when for a station the per-class service times	Users	25.0000	19.0000
are different, the results are correct ONLY IF its scheduling discipline is assumed Processor Sharing (PS) and not FCFS (See BCMP Theorem).			

Στην συνέχεια ορίσαμε Preference Station για τις 2 κατηγορίες.



Και στο τέλος ορίσαμε στην what-if analysis ότι θέλουμε να μελετήσουμε με παράμετρο το **Population Mix**.

Ένα ενδιαφέρον ερώτημα είναι το πόσα βήματα θα πρέπει να χρησιμοποιήσουμε για να έχουμε ακριβής αποτέλεσμα στη μελέτη μας.

Είναι προφανές ότι όσο μικρότερο είναι το n τόσο πιο γρήγορη θα είναι και η προσομοίωση αλλά όσο πιο μεγάλο είναι το n τόσο πιο ακριβής είναι η προσομοίωση, αυτό είναι ένα trade off που πρέπει να αντιμετωπίσουμε.

Το ερώτημα αυτό μπορεί να απαντηθεί μόνο αν ορίσουμε το πόσο ακρίβεια θέλουμε να έχουμε στην προσομοίωση μας. Σε αυτή την άσκηση θα μελετήσουμε το πρόβλημα με τη μέγιστη δυνατή ακρίβεια δηλαδή για n=479.

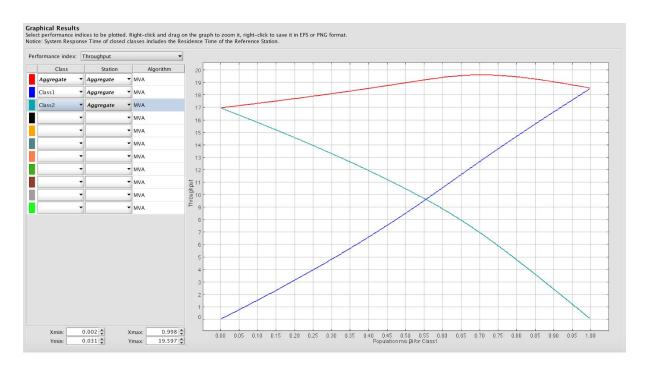
Ωστόσο, τρέξαμε την προσομοίωση και για την default τιμή των n=11 βημάτων και είδαμε ότι το αποτέλεσμα είναι σχεδόν ίδιο - διαφοροποίηση κατα 2 μόλις άτομα στο βέλτιστο μειγμα πληθυσμού όσον αφορά την συνολική ρυθμαπόδοση του συστήματος.



Τέλος, αφήσαμε τον default algorithm ανάλυσης, τον MVA.

Μετά τους υπολογισμούς λοιπόν προέκυψαν τα ακόλουθα αποτελέσματα:

Αρχικά όσον αφορά το μείγμα πληθυσμού σε σχέση με το throughput βλέπουμε την εξής γραφική:



Η κλάση 1 είναι οι χρήστες εγγραφής και χρωματίζεται με μπλε χρώμα. Η κλάση 2 είναι οι χρήστες ανανεώσεις και χρωματίζεται με πράσινο χρώμα. Ενώ με κόκκινο χρώμα χρωματίζεται η συνολική ρυθμαπόδοση του συστήματος. Βλέπουμε λοιπόν ότι η συνολική ρυθμαπόδοση για την κλάση 1 βελτιώνεται όσο μεγαλώνει το πλήθος του πληθυσμού της και μειώνεται το πλήθος του πληθυσμού της κλάσης 2. Ενώ, η συνολική ρυθμαπόδοση για την κλάση 2 χειροτερεύει όσο μικραίνει ο πληθυσμός της

και μεγαλώνει ο πληθυσμός της κλάσης 2.

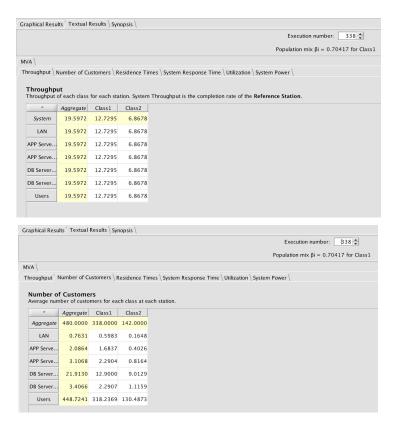
Η συνολική ρυθμαπόδοση αρχικά βελτιώνει όσο αυξάνεται ο πληθυσμός της κλάσης 1 και μειώνεται ο πληθυσμός της κλασης 2 αλλά μετά από ένα σημείο, ένα ακρότατο, χειροτερεύει όσο αυξάνεται ο πληθυσμός της κλάσης 1 και μειώνεται ο πληθυσμός της κλασης 2. Εμείς καλούμαστε να βρούμε αυτό το τοπικό ακρότατο, η γραφική χάνει ακρίβεια αλλά βλέπουμε ότι το ακρότατο είναι κοντά στο 0.7 δηλαδή για πληθυσμό κλάσης 1 ίσο με 0.7*480=336 και για πληθισμός της κλάσης 2 ίσο με 144.

Ωστόσο επειδή θέλουμε ακρίβεια, γι' αυτό και βάλαμε το μέγιστο δυνατό n θα ελέγξουμε αυτή την τιμή από το textual result.

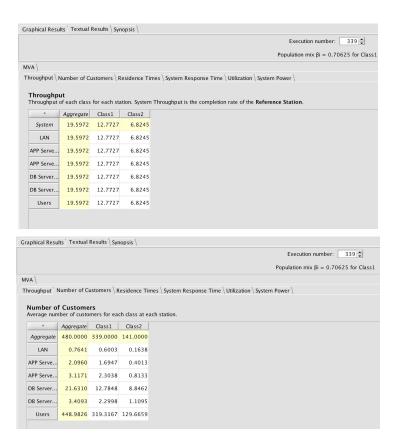
Βλέπουμε ότι η βέλτιστη ρυθμαπόδοση επιτυγχάνεται με δύο μείγματα πληθυσμού, δηλαδή η βέλτιστη συνολική ρυθμαπόδοση είναι ίση με 19.5972 και επιτυγχάνεται με

- β1= 0.70417 δηλαδή για το μείγμα με πληθυσμός κλάσης 1 ίσο με 338 και με πληθυσμός κλάσης 2 ίσο με 142.
- και β2 = 0.70625 δηλαδή για το μείγμα με πληθυσμός κλάσης 1 ίσο με 339 και με πληθυσμός κλάσης 2 ίσο με 141.

Το πρώτο συμπέρασμα εξάγεται από τα εξείς αποτελέσματα:



Ενώ το δεύτερο συμπέρασμα εξάγεται από τα εξείς αποτελέσματα:



Έτσι, επιλέγουμε τυχαία έναν από τους δύο ως βέλτιστο συνδυασμό.

Άρα, η μέγιστη συνολική ρυθμαπόδοση επιτυγχάνεται για: 338 χρήστες εγγραφής (κλάσης 1) και 142 χρήστες ανάγνωσης (κλάσης 2).

Χρησιμοποιείτε αυτό το μείγμα για τα επόμενα ερωτηματα.

Ο ρυθμός απόδοσης του συστήματος συνολικά και για καθεμία από τις κατηγορίες πελατών είναι:

Throughp Throughput	ut of each class	for each sta	tion. System	ı Through _l
*	Aggregate	Class1	Class2	
System	19.5972	12.7295	6.8678	

Ο χρόνος απόκρισης του συστήματος συνολικά και για καθεμία από τις κατηγορίες πελατών είναι:

,	Tĥe global a A: This value B: This value	of System R	he "System R esponse Tim esponse Tim	ne includes ne does NO	me" and is obtained weighting the aggregated values by the relative per-class throughput. the Residence Time of the Reference Station. T include the Residence Time of the Reference Station. Iways coincides with the arrival process. Thus the B values are not computed.
	*	Aggregate	Class1	Class2	
	Α	24.4933	26.5525	20.6763	
	В	1.5959	1.5525	1.6763	

Εμείς επικεντρώνουν την προσοχή μας στην περίπτωση Α, όπου συνυπολογίζεται και ο χρόνος σκέψης και δράσης των χρηστών.

Υπενθυμίζουμε ότι ο χρόνος σκέψης και δράσης της κλάσης 1 είναι 25sec ενώ της κλάσης 2 19sec.

Άρα, συνολικά η ένας χρήστης της κλάσης 1 πρέπει να περιμένει 1.5525 sec για να εκτελεσει την εργασια του το συστημα, ενώ ένας χρήστης της κλάσης 1 πρέπει να περιμένει 1.6763 sec για να εκτελέσει την εργασια του το συστημα.

Χρόνος αρκετά μεγάλος επομένως θα πρέπει να προσπαθήσουμε να βελτιώσουμε το σύστημα μας.

Ο βαθμός χρησιμοποίησης των σταθμών του συστήματος είναι ο εξής:

Utilization	1		
Utilization of	a customer	class at the	selected stat
*	Aggregate	Class1	Class2
LAN	0.7631	0.5983	0.1648
APP Serve	0.6779	0.5474	0.1305
APP Serve	0.7593	0.5601	0.1992
DB Server	0.9780	0.5728	0.4052
DB Server	0.7760	0.5219	0.2541
Users	448.7241	318.2369	130.4873

όπου παρουσιάζονται με την σειρά η χρησιμοποίηση των σταθμών(τόσο συνολικά όσο και ανα κλάση): LAN, APP Server CPU, APP Server Disk, DB Server CPU, DB Server Disk και Users.

Προσοχή στην ανάγνωση του πίνακα όπου η συνολική χρησιμοποιήση του σταθμού lan πρέπει να μεταφραστεί ώς 76.31%. Το ίδιο ισχύει και για τους υπόλοιπους σταθμούς.

Γενικός Σχολιασμός Αποτελεσμάτων Α Ερωτήματος

Θα μπορούσε με μια πρώτη ανάγνωση να έκανε εντύπωση το αποτέλεσμα της μιξης πληθυσμών που βελτιστοποιεί την ρυθμαπόδοση δεδομένου ότι σχεδόν σε όλους του σταθμούς του δικτύου (εκτός από την cpu του db server) έχουμε ότι οι χρήστες κλάσης 2, δηλαδή η χρήστες ανανέωσης έχουν λιγότερο station time.

Ωστόσο, αυτό το αποτέλεσμα του περίπου 70-30 καθορίζεται από το γεγονός ότι έχουν διαφορετικούς term χρόνους, δηλαδή οι χρήστες κλάσης 1 έχουν μεγαλύτερο term χρόνο όπου δεν χρησιμοποιούν το δίκτυο, δηλαδή τους σταθμους του συστήματος, αλλά χρησιμοποιούν το UI κομμάτι του συστήματος. Γι αυτό και η ρυθμαπόδοση μελετάται με βάση το term της κάθε κλάσης.

Παρατηρούμε από την συνολική χρησιμοποίηση των σταθμών (του συστήματος- εκτος του χρονου σκέψης των χρηστών) την πιο μεγάλη τιμή παρουσιάζει ο **DB Server CPU** όπότε εκεί εντοπίζεται η **στένωση** του συστήματος!

Βέβαια το αποτέλεσμα αυτό δεν θα μπορούσε να μας προκαλέσει έκπληξη καθώς παρατηρούμε ότι και για τις δύο κλάσεις στην ΚΜΕ του db server σπαταλιέται ο περισσότερος χρόνος, από κάθε άλλο σταθμό του συστήματος (προφανώς δεν μετράμε το term σε αυτή την ανάλυση).

Επίσης το γεγονός ότι η κμε του db server χρησιμοποιείται κατά (περίπου) 98% θα δημιουργεί πολλά προβλήματα στην λειτουργία της όσο και στις ανάγκες ψυξης της.

Σχετικά με τον χρόνο αναμονής που είναι περίπου 1.5sec για την ανταπόκριση του συστήματος είναι κατανοητό ότι δεν θέλουμε οι χρήστες να περιμένουν τόσο πολύ για το σύστημα να ανταποκριθεί αλλά πρέπει να βελτιώσουμε την υποδομή του.

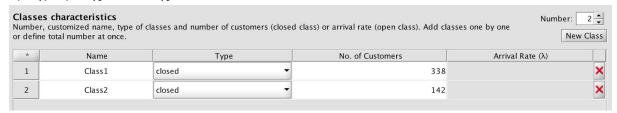
Ερώτημα Β

Προκειμένου να βελτιωθεί η επίδοση του συστήματος μελετάται η αντικατάσταση της CPU του εξυπηρετητή βάσης δεδομένων με άλλη υψηλότερων επιδόσεων. Σημειώνεται ότι, λόγω ειδικών απαιτήσεων ακεραιότητας και ασφάλειας, η νέα CPU θα λειτουργεί σε κατάσταση πολυεπεξεργασίας και θα υφίσταται πρόσθετη υπολογιστική επιβάρυνση. Συνεπώς, η CPU του εξυπηρετητή βάσης δεδομένων θα έχει ρυθμό εξυπηρέτησης εξαρτώμενο από το φορτίο (LD). Υπολογίζεται ότι η απαίτηση εξυπηρέτησης Dij(k), i=5, j=1, 2, όταν στον σταθμό βρίσκονται k εργασίες ανεξαρτήτως κατηγορίας, θα δίνεται από τη σχέση:

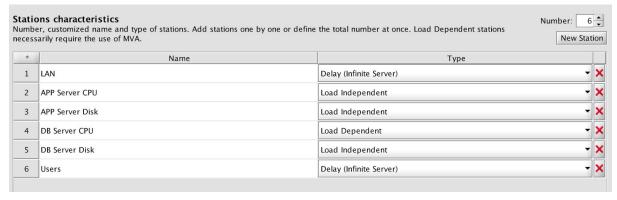
Dij(k) = Dij / (a + b×k) όπου a=5, b=0,01 και Dij η απαίτηση εξυπηρέτησης που αντιστοιχεί στη CPU του Db Server, όπως δίνεται στον Πίνακα για την περίπτωση σταθμού Ll. Ζητείται να υπολογιστούν οι δείκτες επίδοσης όπως στο ερώτημα (α) λαμβάνοντας υπόψη την αντικατάσταση της CPU του Db Server. Υπάρχει βελτίωση σε σχέση με την επίδοση του (α);

Παρατηρήσαμε στο προηγούμενο ερώτημα ότι η στένωση του συστήματος εντοπίζεται στην cpu του db server και τώρα προσπαθούμε να την βελτιώσουμε. Αρχικά, ορίζουμε τον πληθυσμό του συστήματος σύμφωνα με το αποτέλεσμα της

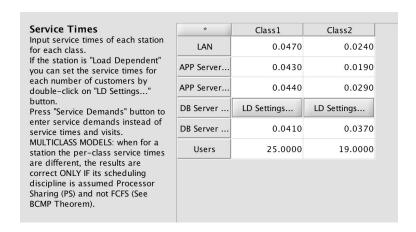
προηγούμενης ανάλυσης.



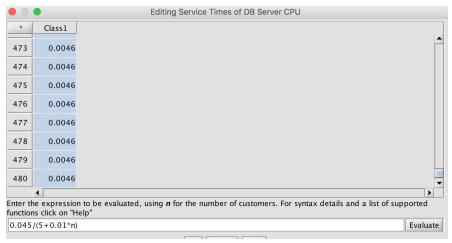
Μετά σειρά έχει η δήλωση των σταθμών του συστήματος και ο χαρακτηρισμός τους ως Load Independent και Delay για τους σταθμούς αναμονής και καθυστέρησης αντίστοιχα, ωστόσο τώρα ρυθμίζουμε ώς load dependent τον σταθμό DB Server CPU.



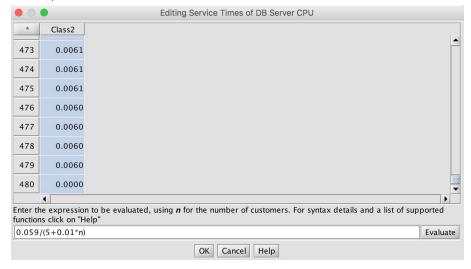
Στη συνέχεια ορίζουμε τις απαιτήσεις σε μέσο χρόνο χρήστη,δηλαδή σε μέσο χρόνο σκέψης, και σε μέση συνολική απαίτηση εξυπηρέτησης για κάθε σταθμό. Οι τιμές που φαίνονται στον πίνακα είναι της τάξεως των sec. Ωστόσο για τον DB Server CPU συνεχίζουμε την ανάλυση και εφαρμόζουμε τον τύπο της εκφώνησης.



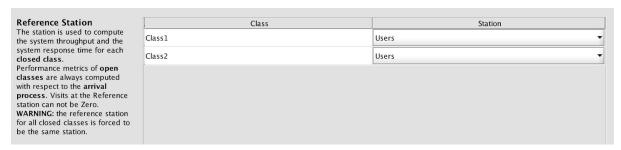
Επιλέγοντας το LD Setting για την κλάση 1 βάζουμε τον εξής τυπο επιλέγοντας όλα τα n και πατάμε evaluate.



Αντίστοιχα, επιλέγοντας το LD Setting για την κλάση 2 βάζουμε τον εξής τυπο επιλέγοντας όλα τα n και πατάμε evaluate.



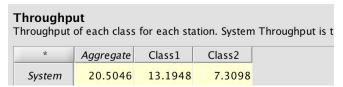
Στην συνέχεια ορίσαμε Preference Station για τις 2 κατηγορίες.



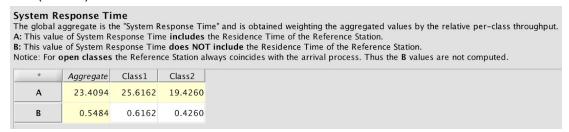
Τέλος, αφήσαμε τον default algorithm ανάλυσης, τον MVA.

Έτσι τρέχοντας την προσομοίωση βλέπουμε ότι:

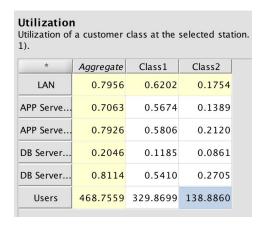
Ο ρυθμός απόδοσης του συστήματος συνολικά και για καθεμία από τις κατηγορίες πελατών είναι:



Ο χρόνος απόκρισης του συστήματος συνολικά και για καθεμία από τις κατηγορίες πελατών είναι(σε sec):



Ο βαθμός χρησιμοποίησης των σταθμών του συστήματος είναι ο εξής:



Προσοχή στην ανάγνωση του πίνακα όπου η συνολική χρησιμοποιήση του σταθμού lan πρέπει να μεταφραστεί ώς 70.63%. Το ίδιο ισχύει και για τους υπόλοιπους σταθμούς.

Γενικός Σχολιασμός Αποτελεσμάτων Β Ερωτήματος

Αρχικά, η συνάρτηση είναι φθίνουσα και έτσι τα station time για την cpu του db server για την κλάση 1 κυμαίνεται μεταξύ των τιμών: 9msec - 4.6msec

Ενώ για τη κλάση 2 κυμαίνονται μεταξύ των τιμών: 11.8msec-6 msec

Οπότε αρχικά παρατηρούμε μια μεγάλη βελτίωση της επίδοσης της cpu.

Ας συγκρίνουμε τώρα τα αποτελέσματα ως προς τις τρεις μετρικές:

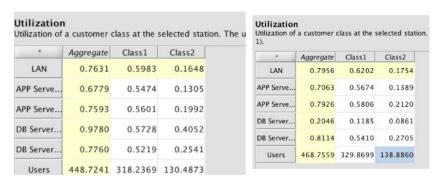
Όπου αριστερά είναι το αποτέλεσμα του ερωτήματος Α και δεξιά το αποτέλεσμα του ερωτήματος Β.



Αρχικά παρατηρούμε βελτίωση της ρυθμαπόδοσης, τόσο της συνολικής όσο και της ρυθμαπόδοσης της κάθε κλάσης και οφείλεται στην βελτίωση της επίδοσης της cpu της db. Η ρυθμαπόδοση ωστόσο δεν είναι τόσο εντυπωσιακή όσο η εντυπωσιακή βελτίωση της βελτίωσης της cpu του db server καθώς οι υπόλοιποι σταθμοι και οι χρόνοι χρήσης παραμένουν σταθερή.

globa nis val nis val	Response Tir I aggregate is th Iue of System Re ue of System Re r open classes	ne "System R esponse Tim esponse Tim	e includes the does NOT i	The global A: This val B: This val	Response Ti aggregate is the ue of System Re ue of System Re open classes	ne "System R esponse Tim esponse Tim	e does NO
				*	Aggregate	Class1	Class2
*	Aggregate	Class1	Class2	*	33 3		
				* A	Aggregate 23.4094	Class1 25.6162	Class2
W	Aggregate	Class1	Class2	*	33 3		

Παρατηρούμε εντυπωσιακή βελτίωση στον χρόνο του συστήματός, όπου πριν ο χρήστης περίμενε περίπου 1.5 sec για το σύστημα να βρεί την απάντηση που αναζητούσε τώρα θέλουμε μόλις 0.5 sec, χρόνος ανεκτός για την δεδομένη εφαρμογή.



Παρατηρούμε ότι αυξήθηκε η χρησιμοποίηση όλων τον σταθμών για όλες τις κατηγορίες και για τις δυο κλάσεις εκτός από την cpu του db server όπου εντοπίζεται μεγάλη πτώση της χρησιμοποιησης της και σε αυτό συμβαίνει εξαιτίας της πολύ πιο καλης station time της. Επίσης έτσι αντιμετωπίσαμε και το πρόβλημα της υπερβολικά μεγάλης χρήσης της cpu του σερβερ της βάσης δεδομένων.

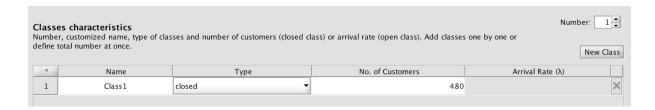
Οπότε συνολικά και ειδικά το σύστημα μας τρέχει πολύ καλύτερα.

Ερώτημα Γ

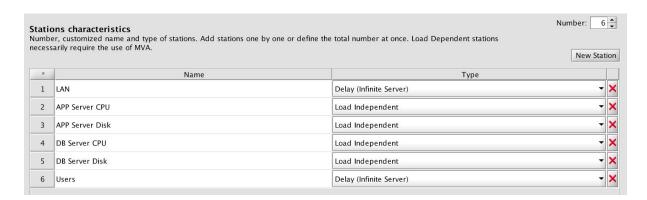
Προκειμένου να απλουστευθεί η ανάλυση, θα ορίσουμε με τα αρχικά δεδομένα αντίστοιχο συγκεντρωτικό μοντέλο μιας κατηγορίας, η οποία θα αποτελεί συνένωση των δύο κατηγοριών. Η τιμή της απαίτησης εξυπηρέτησης σε κάθε σταθμό του μοντέλου μιας κατηγορίας θα τεθεί ίση με τον μέσο όρο των επιμέρους τιμών σταθμισμένων με τον ρυθμό απόδοσης της αντίστοιχης κατηγορίας, όπως υπολογίστηκε στο (α). (Η CPU του Db Server θα ληφθεί ως σταθμός LI, όπως αρχικά.) Να υπολογιστούν οι δείκτες επίδοσης του μοντέλου μιας κατηγορίας και να συγκριθούν με τα αποτελέσματα του (α).

Σε αυτό το σημείο θα προσπαθήσουμε να κάνουμε μια απλούστευση στο μοντέλο μας χρησιμοποιώντας μόνο μια κλάση πληθυσμού η οποία θα συνδυάζει κατάλληλα τις δυο επιμέρους κλάσης του συστήματος μας.

Έτσι αυτή η εννοιά κλάση θα πρέπει να έχει πληθυσμό ίσο με το συνολικό πληθυσμό του συστηματος δηλαδή 480, όπως φαίνεται και στην είσοδο της προσομοίωσης παρακάτω:



Μετά σειρά έχει η δήλωση των σταθμών του συστήματος και ο χαρακτηρισμός τους ως Load Independent και Delay για τους σταθμούς αναμονής και καθυστέρησης αντίστοιχα. Τα ορίζουμε όπως ακριβώς στο πρώτο ερώτημα.



Στη συνέχεια ορίσαμε τις απαιτήσεις σε μέσο χρόνο χρήστη,δηλαδή σε μέσο χρόνο σκέψης, και σε μέση συνολική απαίτηση εξυπηρέτησης για κάθε σταθμό.

Η διαδικασία που ακολουθήθηκε για να βγουν αυτές οι τιμές είναι με χρήση του σταθμισμένου μέσου όρου όπου το w_i είναι το throughput της κλάσης i (i=1,2) όπως αυτό βρέθηκε στο ερώτημα A, x_i είναι το service time ενός συγκεκριμένου station για τη κλάση i, όπως ακριβώς δίνεται στην εκφώνηση. Έτσι λοιπόν ο σταθμισμένος μ.ο δίνεται από τη σχέση:

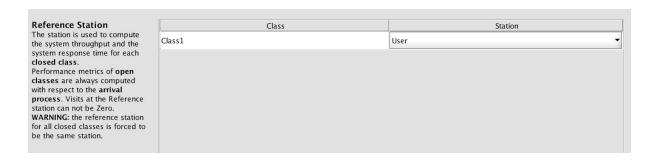
$$\overline{x} = \frac{\sum\limits_{i=1}^{2} w_i \cdot x_i}{\sum\limits_{i=1}^{2} w_i} = \frac{w_1 \cdot x_1 + w_2 \cdot x_2}{w_1 + w_2}$$

π.χ για το πρώτο κελί έχουμε ότι: $\overline{x}_{\alpha} = \frac{47 \cdot 12.7295 + 24 \cdot 6.8678}{19.469} = 39.196$ msec Αντίστοιχα και για τα υπόλοιπα κελιά.

Έτσι, συνολικά έχουμε τον εξής πίνακα, όπου οι τιμές που φαίνονται στον πίνακα είναι της τάξεως των sec:

Service Times	*	Class1
nput service times of each station or each class.	LAN	0.0392
If the station is "Load Dependent" you can set the service times for	APP Server	0.0348
each number of customers by double-click on "LD Settings"	APP Server	0.0390
button. Press "Service Demands" button to	DB Server	0.0502
enter service demands instead of service times and visits.	DB Server	0.0399
MULTICLASS MODELS: when for a station the per-class service times	User	23.0482
are different, the results are correct ONLY IF its scheduling discipline is assumed Processor Sharing (PS) and not FCFS (See BCMP Theorem).		

Στην συνέχεια ορίσαμε Preference Station για την μοναδική κατηγορία μας:



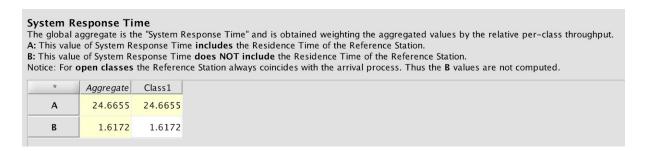
Τέλος, αφήσαμε τον default algorithm ανάλυσης, τον MVA.

Έτσι βλέπουμε ότι:

Ο ρυθμός απόδοσης του συστήματος είναι:

Throughp Throughput		for each sta	ation.
*	Aggregate	Class1	
System	19.4604	19.4604	

Ο χρόνος απόκρισης του συστήματος είναι:



Ο βαθμός χρησιμοποίησης των σταθμών του συστήματος είναι ο εξής:

*	Aggregate	Class1
LAN	0.7628	0.7628
APP Serve	0.6776	0.6776
APP Serve	0.7588	0.7588
DB Server	0.9776	0.9776
DB Server	0.7757	0.7757

Γενικός Σχολιασμός Αποτελεσμάτων Γ Ερωτήματος

Αρχικά ας σχολιάσουμε ότι δεν πηραμε τον κλασσικό μέσο όρο αλλά τον σταθμισμένο ως προς την ρυθμαπόδοση καθώς θέλαμε να έχουμε ένα δίκαιο σύστημα που προσομοιάζει όσο πιο πολύ το αρχικό-αναλυτικό σύστημα. Έτσι του δίνουμε station time σταθμισμένο ως προς την συνολικη ρυθμαπόδοση των κλάσεων, θα μπορούσαμε να ήμασταν ακόμη πιο συγκεκριμένη και με πιο καλή προσέγγιση δίνοντας όχι την συνολική ρυθμαπόδοση της κλάσης αλλά την ρυθμαπόδοση της κλάσης επι του σταθμού. Ωστόσο επειδή η συνολική ρυθμαπόδοση ανάγεται από τις επιμέρους είμαστε με μια μεγαλύτερη απόκλιση κοντά στο αρχικό μοντέλο.

Ας συγκρίνουμε το throughput:

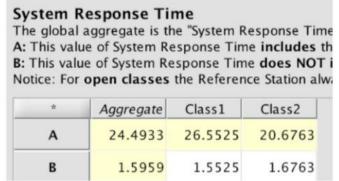
Throughp Throughput	out of each class	for each sta	tion. System	Through
*	Aggregate	Class1	Class2	
System	19.5972	12.7295	6.8678	

hroughp hroughput	of each class	for each stat	ion
*	Aggregate	Class1	
System	19.4604	19.4604	

Αριστερά είναι το throughput στο ερώτημα Α και δεξια στο ερώτημα Β, όπου έχει γίνει η απλουστευση του συστήματος.

Νόημα έχει να συγκρίνουμε μόνο το κελί aggregate στο οποίο βλέπουμε ότι έχουμε σχεδόν ίδια τιμή σε μεγάλο βαθμό, το αν αυτή η απόκλιση γίνει ανεκτή ή όχι εξαρτάται από την ακρίβεια που θέλουμε για τη μελέτη μας.

Ας συγκρίνουμε τον χρόνο απόκρισης του συστήματος:



The global A: This val B: This val	Response Til I aggregate is thue of System Re ue of System Re	ne "System Re esponse Time esponse Time
Notice: For	open classes	Class1
	Aggregate	Class1
Α	24.6655	24.6655

Αριστερά είναι ο χρόνος απόκρισης στο ερώτημα Α και δεξια στο ερώτημα Β, όπου έχει γίνει η απλουστευση του συστήματος.

Νόημα έχει να συγκρίνουμε μόνο τα κελιά aggregate στα οποία βλέπουμε ότι έχουμε σχεδόν ίδιες τιμές σε μεγάλο βαθμό, το αν αυτή η απόκλιση γίνει ανεκτή ή όχι εξαρτάται από την ακρίβεια που θέλουμε για τη μελέτη μας.

Ας συγκρίνουμε τον βαθμό χρησιμοποίησης των σταθμών:

*	Aggregate	Class1	Class2
LAN	0.7631	0.5983	0.1648
APP Serve	0.6779	0.5474	0.1305
APP Serve	0.7593	0.5601	0.1992
DB Server	0.9780	0.5728	0.4052
DB Server	0.7760	0.5219	0.2541
Users	448.7241	318.2369	130.4873

*	Aggregate	Class1
LAN	0.7628	0.7628
APP Serve	0.6776	0.6776
APP Serve	0.7588	0.7588
DB Server	0.9776	0.9776
DB Server	0.7757	0.7757
User	448.5277	448.5277

Αριστερά είναι ο χρόνος απόκρισης στο ερώτημα Α και δεξια στο ερώτημα Β, όπου έχει γίνει η απλουστευση του συστήματος.

Νόημα έχει να συγκρίνουμε μόνο τα κελιά aggregate στα οποία βλέπουμε ότι έχουμε σχεδόν ίδιες τιμές σε μεγάλο βαθμό, το αν αυτή η απόκλιση γίνει ανεκτή ή όχι εξαρτάται από την ακρίβεια που θέλουμε για τη μελέτη μας.

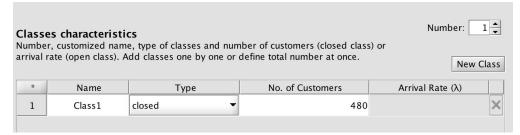
Παρατηρούμε από την συνολική χρησιμοποίηση των σταθμών (του συστήματος- εκτος του χρονου σκέψης των χρηστών) την πιο μεγάλη τιμή παρουσιάζει ο DB Server CPU όπότε εκεί εντοπίζεται η στένωση του συστήματος!

Την ίδια παρατήρηση κάναμε και στο ερώτημα Α, όπου δεν είχαμε θέσει απλοποιήσεις στο σύστημα.

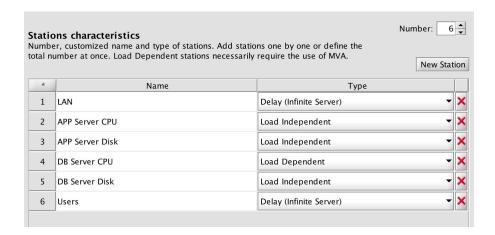
Ερώτημα Δ

Θεωρούμε το μοντέλο μιας κατηγορίας που ορίστηκε στο (γ) υποθέτοντας ότι η CPU του Db Server αντικαθίσταται από σταθμό LD, κατ' αναλογία με το ερώτημα (β). Στο μοντέλο μιας κατηγορίας, η απαίτηση εξυπηρέτησης Di(k), i=5, όταν στον σταθμό βρίσκονται k εργασίες, θα δίνεται από τη σχέση: Di(k) = Di / (a + b×k) όπου a=5, b=0,01 και Di η απαίτηση εξυπηρέτησης που αντιστοιχεί στη CPU του Db Server, όπως υπολογίστηκε στο (γ) για την περίπτωση σταθμού LI. Ζητείται να υπολογιστούν οι δείκτες επίδοσης του μοντέλου μιας κατηγορίας λαμβάνοντας υπόψη την αντικατάσταση της CPU. Υπάρχει βελτίωση σε σχέση με το ερώτημα (γ);

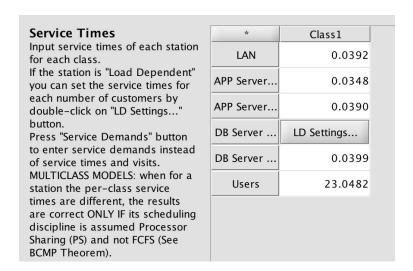
Παρατηρήσαμε στο προηγούμενο ερώτημα ότι η στένωση του συστήματος εντοπίζεται στην cpu του db server και τώρα προσπαθούμε να την βελτιώσουμε. Αρχικά, ορίζουμε τον πληθυσμό του συστήματος:



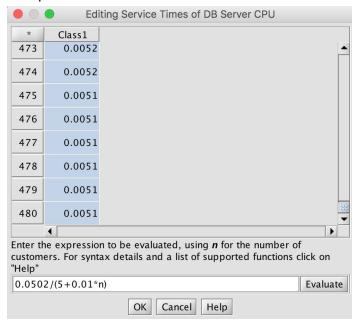
Μετά σειρά έχει η δήλωση των σταθμών του συστήματος και ο χαρακτηρισμός τους ως Load Independent και Delay για τους σταθμούς αναμονής και καθυστέρησης αντίστοιχα, ωστόσο τώρα ρυθμίζουμε ώς load dependent τον σταθμό DB Server CPU.



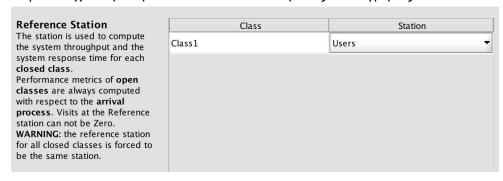
Στη συνέχεια ορίζουμε τις απαιτήσεις σε μέσο χρόνο χρήστη,δηλαδή σε μέσο χρόνο σκέψης, και σε μέση συνολική απαίτηση εξυπηρέτησης για κάθε σταθμό. Οι τιμές που φαίνονται στον πίνακα είναι της τάξεως των sec. Ωστόσο για τον DB Server CPU συνεχίζουμε την ανάλυση και εφαρμόζουμε τον τύπο της εκφώνησης.



Επιλέγοντας το LD Setting για την κλάση 1 βάζουμε τον εξής τυπο επιλέγοντας όλα τα n και πατάμε evaluate.



Στην συνέχεια ορίσαμε Preference Station για τις 2 κατηγορίες.



Τέλος, αφήσαμε τον default algorithm ανάλυσης, τον MVA.

Έτσι βλέπουμε ότι:

Ο ρυθμός απόδοσης του συστήματος είναι:

Throughp Throughput		s for each sta	ation.
*	Aggregate	Class1	
System	20.3367	20.3367	

Ο χρόνος απόκρισης του συστήματος είναι:

System Response TimeThe 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	Class1	
Α	23.6027	23.6027	
В	0.5545	0.5545	

Ο βαθμός χρησιμοποίησης των σταθμών του συστήματος είναι ο εξής:

Utilization

Utilization of a customer class at the customers in the station (it may be gr

*	Aggregate	Class1
LAN	0.5497	0.5497
APP Serve	0.7077	0.7077
APP Serve	0.7931	0.7931
DB Server	0.2037	0.2037
DB Server	0.8114	0.8114
Users	1.0000	1.0000

Γενικός Σχολιασμός Αποτελεσμάτων Δ Ερωτήματος

Αρχικά, η συνάρτηση είναι φθίνουσα και έτσι τα station time για την cpu του db server κυμαίνεται μεταξύ των τιμών: 10msec - 5.1msec

Οπότε αρχικά παρατηρούμε μια μεγάλη βελτίωση της επίδοσης της cpu.

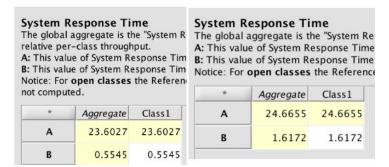
Ας συγκρίνουμε τώρα τα αποτελέσματα ως προς τις τρεις μετρικές: Όπου αριστερά είναι το αποτέλεσμα του ερωτήματος Δ και δεξιά το αποτέλεσμα του ερωτήματος Γ.



Αρχικά παρατηρούμε βελτίωση της ρυθμαπόδοσης, τόσο της συνολικής όσο και της ρυθμαπόδοσης της κάθε κλάσης και οφείλεται στην βελτίωση της επίδοσης της cpu της db. Η ρυθμαπόδοση ωστόσο δεν είναι τόσο εντυπωσιακή όσο η εντυπωσιακή βελτίωση της βελτίωσης της cpu του db server καθώς οι υπόλοιποι σταθμοι και οι χρόνοι χρήσης παραμένουν σταθερή.

Άρα παρατηρούμε όντως επίδοση τόσο στο απλοποιημένο όσο και στο πιο ρεαλιστικό μοντέλο.

Αριστερά είναι το αποτέλεσμα του ερωτήματος Δ και δεξιά το αποτέλεσμα του ερωτήματος Γ.



Παρατηρούμε εντυπωσιακή βελτίωση στον χρόνο του συστήματός, όπου πριν ο χρήστης περίμενε περίπου 1.6 sec για το σύστημα να βρεί την απάντηση που αναζητούσε τώρα θέλουμε μόλις 0.55 sec, χρόνος ανεκτός για την δεδομένη εφαρμογή.

Άρα παρατηρούμε όντως επίδοση τόσο στο απλοποιημένο όσο και στο πιο ρεαλιστικό μοντέλο.

Αριστερά είναι το αποτέλεσμα του ερωτήματος Δ και δεξιά το αποτέλεσμα του ερωτήματος Γ.

	a customer c the station (i		Utilization Utilization of		class at the
*	Aggregate	Class1	*	Aggregate	Class1
LAN	0.5497	0.5497	LAN	0.7628	0.7628
APP Serve	0.7077	0.7077	APP Serve	0.6776	0.6776
APP Serve	0.7931	0.7931	APP Serve	0.7588	0.7588
DB Server	0.2037	0.2037	DB Server	0.9776	0.9776
DB Server	0.8114	0.8114	DB Server	0.7757	0.7757
Users	1.0000	1.0000	User	448.5277	448.5277

Παρατηρούμε ότι αυξήθηκε η χρησιμοποίηση όλων τον σταθμών για όλες τις κατηγορίες και για τις δυο κλάσεις εκτός από την cpu του db server όπου εντοπίζεται μεγάλη πτώση της χρησιμοποιησης της και σε αυτό συμβαίνει εξαιτίας της πολύ πιο καλης station time της. Επίσης έτσι αντιμετωπίσαμε και το πρόβλημα της υπερβολικά μεγάλης χρήσης της cpu του σερβερ της βάσης δεδομένων.

Οπότε συνολικά και ειδικά το σύστημα μας τρέχει πολύ καλύτερα και **υπάρχει βελτίωση**. Επίσης παρατηρούμε ότι εξάγουμε τα ίδια ακριβώς συμπεράσματα στο ρεαλιστικο και στο απλοποιημένο μοντέλο. Επομένως η απλοποίηση ήταν εύστοχη.

Αξίζει να σημειωθεί ότι το γεγονός ότι το utilization για τους users (term) βγαινει ίσο με την μονάδα το οποίο είναι λάθος και προκύπτει από πιθανό **bug** του συστήματος. Κανονικά πρέπει να είναι ίσο με το μέσο αριθμό εργασιών!