ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

Μελέτη Οικονομικών Αλγορίθμων για Κατανομή Φόρτου Εργασιών και Διαχείριση Δεδομένων σε Κατανεμημένα Συστήματα

Αναστασία Αναστασιάδη

Μεταπτυχιακή Εργασία

Ηράκλειο, Οκτώβριος 1996

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

Μελέτη Οικονομικών Αλγορίθμων για Κατανομή Φόρτου Εργασιών και Διαχείριση Δεδομένων σε Κατανεμημένα Συστήματα

Εργασία που υποβλήθηκε από την
Αναστασία Αναστασιάδη
ως μερική εκπλήρωση των απαιτήσεων
για την απόκτηση
ΜΕΤΑΠΤΥΧΙΑΚΟΥ ΔΙΠΛΩΜΑΤΟΣ ΕΙΔΙΚΕΥΣΗΣ

Συγγραφέας:

Αναστασία Αναστασιάδη
Τμήμα Επιστήμης Υπολογιστών

Εισηγητική Επιτροπή:

Χρήστος Νικολάου, Αναπληρωτής Καθηγητής, Επόπτης

Πάνος Κωνσταντόπουλος, Αναπληρωτής Καθηγητής, Μέλος

Κατερίνα Χούστη, Αναπληρώτρια Καθηγήτρια, Μέλος

Δεκτή:

Πάνος Κωνσταντόπουλος, Αναπληρωτής Καθηγητής
Πρόεδρος Επιτροπής Μεταπτυχιακών Σπουδών

Ηράκλειο, Οκτώβριος 1996

Μελέτη Οικονομικών Αλγορίθμων για Κατανομή Φόρτου Εργασιών και Διαχείριση Δεδομένων σε Κατανεμημένα Συστήματα

Αναστασία Αναστασιάδη

Μεταπτυχιακή Εργασία

Τμήμα Επιστήμης Υπολογιστών
Πανεπιστήμιο Κρήτης

ΠΕΡΙΛΗΨΗ

Οι πρόσφατες εξελίξεις στην τεχνολογία υπολογιστών και δικτύων επιτρέπουν την διασύνδεση μεγάλου αριθμού ετερογενών υπολογιστών και τη δημιουργία μεγάλων συλλογών από υπολογιστικούς και επικοινωνιακούς πόρους. Τα συστήματα αυτά παρουσιάζουν μεγάλη πολυπλοκότητα στην οργάνωση και διαχείριση των πόρων και των υπηρεσιών που διαθέτουν. Η αυξημένη πολυπλοκότητα οφείλεται κυρίως στο μέγεθος (αριθμός συστημάτων, αριθμός χρηστών) και στην ετερογένεια των πόρων και των εφαρμογών και καθιστά τις παραδοσιακές μεθόδους κατανομής πόρων μη αποδοτικές.

Στόχος αυτής της εργασίας είναι να επιδείξει πως ανταγωνιστικά, οικονομικά μοντέλα μπορούν να παρέχουν αποδοτικούς αλγόριθμους για την κατανομή πόρων σε ένα κατανεμημένο υπολογιστικό σύστημα περιορίζοντας σημαντικά την πολυπλοκότητα σε σχέση με τις παραδοσιακές μεθόδους.

Εχουν σχεδιαστεί αλγόριθμοι κατανομής φόρτου εργασιών και διαχείρισης δεδομένων που στηρίζονται στην οικονομία τιμών. Καταναλωτές αυτής της οικονομίας θεωρούνται οι δοσοληψίες οι οποίες εκμεταλλεύονταιτο κεφάλαιο που διαθέτουν για την αγορα πόρων (υπολογιστικός χρόνος, εύρος επικοινωνίας, δεδομένα). Στόχος των δοσοληψιών είναι η ικανοποίηση των αναγκών τους με τη μικρότερη δυνατή χρέωση. Προμηθευτές της

οικονομίας θεωρούνται οι κόμβοι του συστήματος οι οποίοι εμπορεύονται τα τοπικά τους αγαθά και στοχεύουν στη μεγιστοποίηση των κερδών τους.

Η μελέτη της συμπεριφοράς των οικονομικών αλγορίθμων κατανομής φόρτου εργασιών και διαχείρισης δεδομένων έγινε με την διεξαγωγή πλήθους πειραμάτων προσομοίωσης στον προσομοιωτή TPsim. Τα πειράματα πιστοποίησαν οτι οι οικονομικές μέθοδοι επιλύουν αποδοτικά το πρόβλημα της κατανομής πόρων σε κάθε κατανεμημένο συστήμα και περιορίζουν την πολυπλοκότητα της διαδικασίας σε σχέση με τις παραδοσιακές μεθόδους.

Επόπτης : Χρήστος Νικολάου
Αναπληρωτής Καθηγητής Επιστήμης Υπολογιστών
Πανεπιστήμιο Κρήτης

A study of Microeconomic Algorithms for Load Balancing and Data Replication in Distributed Computer Systems

Anastasia Anastasiadi

Master of Science Thesis

Department of Computer Science
University of Crete

ABSTRACT

With the recent advances in computer and networking technology thousands of heterogeneous computers can be interconnected to provide a large collection of computing and communication resources. A macroscopic view of these systems reveals the complexity of the organization and management of the resources and services they provide. This complexity arises from size (no of systems, no of users) and heterogeneity of applications and resources and makes the traditional approaches to resource allocation impractical in modern distributed systems.

The goal of this work is to demonstrade how competitive ecomonic models provide efficient algorithms and tools for allocating resources in distributed computer systems and manage to limit the complexity of resource allocation.

In this work we model the distributed system as a competitive society of microeconomic agents (price based economy) and we apply this model to the problem of load balancing and managing distributed replicated data objects. The consumers of the economy are the transactions which are endowed with some wealth (budget). Each transaction is using its budget to purchase resources (CPU time, communication bandwidth, data objects) with the minimum cost. The nodes of the system are the suppliers of the economy. A supplier's sole goal

is to optimize its individual satisfaction (profit) derived from its choice of resource allocation to consumers.

We evaluate the load balancing economy and the data replication economy through a simulation study on TPsim. Our experiments show that the microeconomic algorithms can substantially improve the performance and reduce the complexity of resource allocation relative to traditional approaches.

Supervisor : Christos Nikolaou
Associate Professor of Computer Science
University of Crete

Ευχαριστίες

Στο σημείο αυτό αισθάνομαι την ανάγκη να διατυπώσω τις ευχαριστίες μου σε όλους όσους με τον ένα ή τον άλλο τρόπο με βοήθησαν να φέρω σε πέρας αυτή την εργασία.

Πρώτα απ' όλα ευχαριστώ το Τμήμα Επιστήμης Υπολογιστών του Πανεπιστημίου Κρήτης που με έκανε δεκτή στο μεταπτυχιακό του πρόγραμμα δίνοντας μου την ευκαιρία να αναπτύξω τις ερευνητικές μου προσπάθειες, καθώς και το Ινστιτούτο Πληροφορικής του Ι.Τ.Ε. που με την οικονομική και υλικοτεχνική υποστήριξη που μου παρείχε, στήριξε τις προσπάθειες αυτές.

Ιδιαίτερα ευχαριστώ τον επόπτη μου κ. Χρήστο Νικολάου που μου έδωσε την δυνατότητα να ασχοληθώ με το θέμα της κατανομής πόρων χρησιμοποιώντας οικονομικούς αλγορίθμους και με βοήθησε με τις πολύτιμες υποδείξεις του στην εκπόνηση αυτής της εργασίας.

Ευχαριστώ επίσης τον Jakka Sairamesh για την καθοδήγηση του σε όλη τη διάρκεια της εργασίας μου και για τις εκτενής συζητήσεις που είχαμε και με βοήθησαν στην κατανόηση των οικονομικών μοντέλων.

Θέλω να ευχαριστήσω όλα τα μέλη της ομάδας των Πλειάδων του Ινστιτούτου Πληροφορικής για τη συνεργασία και τη βοήθεια που απλόχερα μου πρόσφεραν όποτε και για ό,τι την είχα ανάγκη. Ειδικότερα θα ήθελα να ευχαριστήσω τον Μανώλη Μαραζάκη για τη βοήθεια του στην εκμάθηση του προσομοιωτή TPsim και τους Σαράντο Καπιδάκη, Σωτήρη Τερζή και David McClarkin για όλες τις εποικοδομητικές συζητήσεις που είχαμε.

Θέλω επίσης να ευχαριστήσω τον Βασίλη Σπιταδάκη και την Ανθή Γιώρτσου για τα εκτενή και ουσιώδη σχόλια τους σχετικά με την αρχική έκδοση του κειμένου εργασίας.

Την Ανθή αλλά και τις Μαρίνα Μωραΐτη και Άντζη Χλαπάνη θα ήθελα ιδιαίτερα να τις ευχαριστήσω γιατί αποτέλεσαν για μένα μια οικογένεια στο Ηράκλειο που με στήριξε στις δύσκολες στιγμές και έκανε ακόμα πιο πολλές και πιο όμορφες τις καλές στιγμές.

Ένα μεγάλο ευχαριστώ πρέπει να πω και στο Χρίστο Χαμηλάκη για την βοήθεια του στην προετοιμασία του κειμένου της εργασίας και της παρουσίασης. Ιδιαίτερα όμως τον ευχαριστώ για την αγάπη που μου έδειξε και την συναισθηματική στήριξη που μου πρόσφερε σε όλη τη διάρκεια της φοιτητικής μου ζωής.

Τέλος θα ήθελα να ευχαριστήσω τους γονείς μου Παναγιώτη και Γεωργία καθώς και την αδερφή μου Ζωή. Η αγάπη και η εμπιστοσύνη τους με έκαναν να πιστέψω στον εαυτό μου, να αντλήσω δύναμη και να πραγματοποιήσω τους στόχους μου. Σε αυτούς αφιερώνω αυτήν την εργασία.

Περιεχόμενα

Περίληψη **i**

Abstract **iii**

Ευχαριστίες **v**

1 Εισαγωγή **1**

1.1 Οικονομικά Μοντέλα 2

1.1.1 Γιατί να χρησιμοποιηθούν οικονομικά μοντέλα; 2

1.1.2 Κοστολόγηση και απόδοση 3

1.1.3 Προτιμήσεις καταναλωτών 3

1.1.4 Ανταγωνιστικά μοντέλα 4

1.1.5 Τι κοστολογείται και ποιος έχει τον έλεγχο 5

1.2 Οργάνωση εργασίας 6

2 Ο Προσομοιωτής TPsim **9**

2.1 Το Μοντέλο Συστήματος 9

2.2 Γλώσσα Προδιαγραφής 11

2.2.1 Περιγραφή Συσκευών Αποθήκευσης 11

2.2.2 Περιγραφή Κόμβων και Δικτύου Κόμβων 12

2.2.3 Περιγραφή Βάσης Δεδομένων 15

2.2.4 Περιγραφή Φόρτου Εξυπηρέτησης 17

2.3 Προσθήκες στο Tpsim για την υποστήριξη οικονομικών αλγορίθμων 18

3 Οικονομία Κατανομής Φόρτου Εργασιών **21**

3.1 Περιγραφή Προβλήματος 21

3.2 Περιγραφή Οικονομικών Αλγορίθμων Δρομολόγησης 24
3.3 Μοντέλα Προσομοίωσης 28
3.3.1 Μοντέλα με μία πηγή φόρτου 28
3.3.2 Μοντέλα με πολλαπλές πηγές φόρτου 31
3.4 Αποτελέσματα 33
3.5 Ανάλυση 41
3.6 Επισκόπηση Σχετικής Βιβλιογραφίας 43

4 Οικονομία Διαχείρισης Δεδομένων **45**
4.1 Περιγραφή Προβλήματος 45
4.2 Οικονομικός αλγόριθμος αντιγραφής δεδομένων και δρομολόγησης δοσοληψιών DRR 47
4.3 Μοντέλα προσομοίωσης 52
4.4 Αποτελέσματα 53
4.5 Επισκόπηση Σχετικής Βιβλιογραφίας 61

5 Συμπεράσματα και Ερευνητικές Κατευθύνσεις **63**
5.1 Σύνοψη Αποτελεσμάτων 63
5.2 Επεκτάσεις και Ερευνητικές Κατευθύνσεις 65

Βιβλιογραφία **67**

Κεφάλαιο 1

Εισαγωγή

Οι πρόσφατες εξελίξεις στην τεχνολογία υπολογιστών και δικτύων επιτρέπουν την διασύνδεση μεγάλου αριθμού ετερογενών υπολογιστών και τη δημιουργία μεγάλων συλλογών από υπολογιστικούς και επικοινωνιακούς πόρους. Τα συστήματα αυτά χρησιμοποιούνται από ένα αυξανόμενο αριθμό χρηστών και παρουσιάζουν μεγάλη πολυπλοκότητα στην οργάνωση και διαχείριση των πόρων και των υπηρεσιών που διαθέτουν. Η πολυπλοκότητα αυτή οφείλεται στο μέγεθος (αριθμός συστημάτων, αριθμός χρηστών) στην ετερογένεια των εφαρμογών και των πόρων. Διάφοροι παράγοντες επαυξάνουν την πολυπλοκότητα της κατανομής πόρων. Σε πολλά κατανεμημένα συστήματα ποικίλοι "οργανισμοί" θεωρούνται ιδιοκτήτες των πόρων. Επίσης, η ικανοποίηση των χρηστών και η απόδοση των εφαρμογών καθορίζεται από την ταυτόχρονη κατανομή διαφόρων πόρων. Για παράδειγμα, ένας πολύμορφος εξυπηρετητής εφαρμογών απαιτεί I/O bandwidth για την ανάκληση δεδομένων, υπολογιστικό χρόνο για την εκτέλεση πρωτοκόλλων εξυπηρέτησης και επικοινωνίας και εύρος δικτύου για την μεταφορά του δεδομένου στους πελάτες του συστήματος. Η απόδοση των εφαρμογών δύναται να επηρεαστεί από την ανταλλαγή αγαθών. Για παράδειγμα, ο πολύμορφος εξυπηρετητής εφαρμογών δύναται να αποδώσει καλύτερα αφήνοντας μνήμη και απαιτώντας μεγαλύτερη υπολογιστική προτεραιότητα. Η διαπραγμάτευση αυτή μπορεί να οδηγήσει σε μικρότερους ενταμιευτές αλλά βελτιώνει την απόδοση εκτέλεσης του πρωτοκόλλου επικοινωνίας. Τέλος, σε ένα μεγάλο κατανεμημένο σύστημα, ο αριθμός των συστημάτων, των χρηστών και των εφαρμογών αλλάζει συνεχώς. Αυτή η αυξημένη πολυπλοκότητα καθιστά τους παραδοσιακούς αλγόριθμους κατανομής πόρων μη αποδοτικούς στα μοντέρνα κατανεμημένα συστήματα. Οι παραδοσιακές

μέθοδοι επιχειρούν να βελτιστοποιήσουν μία μεταβλητή απόδοσης, που αφορά το σύστημα σε όλο του το εύρος. Η βελτιστοποίηση αυτή επιτυγχάνεται, είτε με ένα κεντρικοποιημένο αλγόριθμο με πλήρη πληροφορία, είτε με ένα αποκεντρικοποιημένο αλγόριθμο κοινής συναίνεσης. Η πολυπλοκότητα κατανομής πόρων στα σύγχρονα κατανεμημένα συστήματα καθιστά αδύνατο τον ορισμό μιας αποδεκτής, σε όλο το εύρος του συστήματος, μεταβλητής απόδοσης. Ποιά μεταβλητή απόδοσης είναι ικανή να εκφράσει με επάρκεια τους αντικειμενικούς στόχους απόδοσης ενός πολύμορφου εξυπηρετητή εφαρμογών ή ενός online συστήματος επεξεργασίας δοσοληψιών;

Η πολυπλοκότητα κατανομής πόρων λόγω αποκεντρικοποίησης και ετερογένειας, παρουσιάζεται επίσης και στις ανθρώπινες οικονομίες. Στόχος αυτής της εργασίας είναι να επιδείξει πως ανταγωνιστικά οικονομικά μοντέλα μπορούν να παρέχουν αλγόριθμους για την κατανομή πόρων σε ένα υπολογιστικό σύστημα.

1.1 Οικονομικά Μοντέλα

1.1.1 Γιατί να χρησιμοποιηθούν οικονομικά μοντέλα;

1. **Μείωση πολυπλοκότητας**

   Τα οικονομικά μοντέλα περιορίζουν την πολυπλοκότητα του προβλήματος κατανομής πόρων με αποκεντρικοποίηση του ελέγχου των πόρων. Κάθε κόμβος του κατανεμημένου συστήματος, ελέγχει τους πόρους που διαθέτει.

2. **Αποκεντρικοποίηση**

   Σε μια οικονομία υπάρχουν δύο τύποι πρακτόρων, οι προμηθευτές και οι καταναλωτές. Ενας καταναλωτής επιδιώκει να καλύψει τις ανάγκες του σε πόρους και υπηρεσίες και έτσι να βελτιστοποιήσει τους προσωπικούς του στόχους απόδοσης. Ενας προμηθευτής επιδιώκει να κατανέμει κατάλληλα τα προϊόντα του στους καταναλωτές ώστε να μεγιστοποιήσει τα προσωπικά του κέρδη. Και οι δύο αδιαφορούν για την απόδοση του υπόλοιπου συστήματος. Σε μια οικονομία λοιπόν η αποκεντρικοποίηση επιτυγχάνεται από το γεγονός ότι οι οικονομικοί πράκτορες επιδιώκουν να επιτύχουν τους στόχους τους με ανταγωνιστικό, εγωϊστικό τρόπο. Πλήθος εργασιών εφαρμόζουν το μοντέλο αυτό για να επιτύχουν αποκεντρικοποιημένη κατανομή πόρων σε υπολογιστικά συστήματα [5, 6, 7, 8, 10, 15, 18, 23, 24, 26]

1.1.2 Κοστολόγηση και απόδοση

Στα περισσότερα οικονομικά μοντέλα εισάγονται οι έννοιες του χρήματος και της κοστολόγησης για τον συντονισμό των πρακτόρων της οικονομίας. Κάθε καταναλωτής προικίζεται με ένα ποσό χρημάτων το οποίο χρησιμοποιεί για την αγορά των πόρων που έχει ανάγκη. Κάθε προμηθευτής κατέχει ένα σύνολο πόρων τους οποίους οι καταναλωτές μπορούν να χρησιμοποιήσουν μετά την καταβολή χρηματικού ποσού. Κάθε προμηθευτής καθορίζει την τιμή ενός προϊόντος του (πόρου), με βάση την προσφορά και την ζήτηση που έχει σημειωθεί για το προϊόν αυτό. Το οικονομικό σύστημα εγγυάται την ύπαρξη μιας πραγματοποιήσιμης κατανομής πόρων [17].

1.1.3 Προτιμήσεις καταναλωτών

Εστω οικονομία που αποτελείται από τα αγαθά (πόρους) $r_1, r_2, .., r_M$. Μια ανάθεση των αγαθών αυτών σε ένα πράκτορα του συστήματος είναι ένα διάνυσμα $\chi = \{x_1, x_2, .., x_M\}$ όπου x_i είναι η ποσότητα του αγαθού r_i που έχει ανατεθεί στον πράκτορα αυτό. Κάθε πράκτορας της οικονομίας προτιμά κάποιες αναθέσεις αγαθών, περισσότερο από κάποιες άλλες. Οι προτιμήσεις των πρακτόρων εκφράζονται από την συνάρτηση χρησιμότητας (utility function) [1, 2]. Η συνάρτηση αυτή είναι απεικόνιση από τον χώρο των διανυσμάτων ανάθεσης αγαθών, στο σύνολο των πραγματικών αριθμών. Εστω $U(.)$ μια συνάρτηση χρησιμότητας και x, y δύο διανύσματα ανάθεσης πόρων. Ο πράκτορας P_i προτιμά την ανάθεση x από την y $(x > y)$ όταν $U(x) > U(y)$. Σε ένα απλό υπολογιστικό σύστημα κάθε ανάθεση υπολογιστικού χρόνου και μνήμης καθορίζει, για κάθε εφαρμογή, ένα χρόνο απόκρισης. Αν ο πράκτορας προτιμά τις αναθέσεις που δίνουν μικρότερο χρόνο απόκρισης, η συνάρτηση χρησιμότητας θα ήταν -1 επί τον χρόνο απόκρισης.

Το σύνολο των αναθέσεων που είναι εφικτές κάτω από τον περιορισμό του κεφαλαίου (budget) για κάθε πράκτορα, δίνεται από τον ακόλουθο τύπο:

$$B(\overrightarrow{p}) = \{x : \overrightarrow{p} \times x \leq w\}$$

όπου w είναι το κεφάλαιο του πράκτορα και $\overrightarrow{p}$ είναι το διάνυσμα τιμών που αντιστοιχεί στα αγαθά (η τιμή του αγαθού r_i είναι p_i). Το σύνολο $B(\overrightarrow{p})$ καθορίζει την εφικτή περιοχή. Το σύνολο ζήτησης $\Phi(\overrightarrow{p})$ για κάθε πράκτορα είναι το σύνολο των διανυσμάτων ανάθεσης αγαθών που βρίσκονται μέσα στην εφικτή περιοχή και η συνάρτηση χρησιμότητάς τους

δίνει μέγιστη τιμή.

$$\Phi(\overrightarrow{p}) = \{x : x \in B(\overrightarrow{p}), U(x) \geq U(x'), \forall x' \in B(\overrightarrow{p})\}$$

Ο πράκτορας λοιπόν ζητά την ανάθεση αγαθών με μεγαλύτερη προτίμηση που είναι εφικτή δεδομένου του περιορισμού κεφαλαίου.

1.1.4 Ανταγωνιστικά μοντέλα

Υπάρχουν δυο βασικοί μηχανισμοί ανάθεσης αγαθών στους ανταγωνιστικούς πράκτορες. Ο ένας είναι η οικονομία ανταλλαγής (exchange based economy) και ο άλλος η οικονομία τιμών (price based economy) [1, 2, 13]. Στην οικονομία ανταλλαγής, κάθε πράκτορας αρχικά προικίζεται με ένα ποσοστό αγαθών. Οι πράκτορες της οικονομίας, στη συνέχεια, ανταλλάσσουν αγαθά μέχρι ο οριακός ρυθμός ανταλλαγής των αγαθών να είναι ίδιος για όλους τους πράκτορες. Δύο πράκτορες θα συμφωνήσουν στην ανταλλαγή των αγαθών, αν η ανταλλαγή αυτή ικανοποιεί τις προτιμήσεις και των δύο πρακτόρων (οδηγεί σε διάνυσμα ανάθεσης αγαθών για το οποίο η συνάρτηση χρησιμότητας δίνει μεγαλύτερη τιμή για κάθε πράκτορα). Η "pareto optimal" ανάθεση επιτυγχάνεται όταν καμία νέα ανταλλαγή δεν μπορεί να οδηγήσει σε αμοιβαία συμφέρουσα ανταλλαγή αγαθών [14].

Σε μια οικονομία τιμών τα αγαθά κοστολογούνται με βάση τη ζήτηση, την προσφορά και το κεφάλαιο των πρακτόρων του συστήματος. Οι αναθέσεις αγαθών πραγματοποιούνται με βάση τον ακόλουθο μηχανισμό: Κάθε πράκτορας προικίζεται με ένα ποσό χρημάτων (κεφάλαιο) και υπολογίζει την ζήτησή του σε αγαθά, με την βοήθεια της συνάρτησης χρησιμότητας και τον περιορισμό κεφαλαίου. Η συνολική ζήτηση των πρακτόρων του συστήματος αποστέλλεται στους προμηθευτές, οι οποίοι υπολογίζουν τις νέες τιμές των αγαθών τους. Αν η ζήτηση για ένα αγαθό είναι μεγαλύτερη από την προσφορά, οι προμηθευτές αυξάνουν την τιμή του αγαθού. Αν αντίθετα η προσφορά δεν καλύπτεται από την ζήτηση, η τιμή του αγαθού μειώνεται. Στη συνέχεια οι πράκτορες υπολογίζουν ξανά την ζήτησή τους για αγαθά, βασιζόμενοι στις καινούριες τιμές και αποστέλουν τη συνολική ζήτηση στους προμηθευτές. Η διαδικασία αυτή επαναλαμβάνεται μέχρι η ζήτηση να ισούται με την προσφορά (τιμή ισορροπίας-equilibrium price).

Ο πλειστηριασμός είναι ένας ακόμη μηχανισμός για ανάθεση αγαθών με βάση τις τιμές τους. Υπάρχουν πολλές μορφές πλειστηριασμού όπως ο Sealed Bid Auction, Dutch

Auction και English Auction. Η βασική φιλοσοφία των πλειστηριασμών είναι οτι τα αγαθά παραχωρούνται στον πλειοδότη που εκφωνεί την μεγαλύτερη προσφορά και η τιμή των αγαθών διαμορφώνεται με βάση τις προσφορές.

1.1.5 Τι κοστολογείται και ποιος έχει τον έλεγχο

Σε ένα κατανεμημένο σύστημα, οι καταναλωτές είναι εφαρμογές όπως δοσοληψίες, αιτήσεις μεταφοράς αρχείων (file transfer applications), πολύμορφες τηλεσυνεδριάσεις (multimedia teleconferences) κ.ο.κ. Οι κόμβοι του συστήματος είναι οι βασικοί προμηθευτές οι οποίοι διαχειρίζονται τοπικούς πόρους, όπως υπολογιστικό χρόνο (CPU), κύρια και κρυφή μνήμη (memory, cache), μαγνητικούς δίσκους (disks), εύρος επικοινωνίας (communication bandwidth) κ.α. Στους προμηθευτές του συστήματος περιλαμβάνονται επίσης εξυπηρετητές (servers) λογικών υπηρεσιών όπως οι name servers, network file servers, web servers κ.α.

Οι προμηθευτές κοστολογούν τα αγαθά τους για να ελέγξουν την διανομή τους, ενώ οι καταναλωτές αγοράζουν αγαθά για να ικανοποιήσουν τις υπολογιστικές τους απαιτήσεις. Οι τιμές διαμορφώνονται από τους προμηθευτές με βάση την ζήτηση των αντίστοιχων αγαθών. Οι τιμές των αγαθών και τα κεφάλαια των καταναλωτών είναι φανταστικά. Το κεφάλαιο κάθε καταναλωτή καθορίζεται από τον διαχειριστή του συστήματος και αποτελεί δείκτη προτεραιότητας του συγκεκριμένου καταναλωτή. Οσο μεγαλύτερο είναι το κεφάλαιο τόσο περισσότερα αγαθά, ή ποσοστό ενός αγαθού, μπορεί να αγοράσει ο καταναλωτής. Στην διαμόρφωση ενός οικονομικού πλαισίου εργασίας για παροχή υπηρεσιών σε ένα κατανεμημένο σύστημα υπάρχουν ορισμένοι περιορισμοί-απαιτήσεις:

- Τι πρέπει να κοστολογηθεί; Ποιοι πόροι (ενταμιευτές, υπολογιστικός χρόνος, εύρος επικοινωνίας) καθώς και υπηρεσίες που βασίζονται σε παραμέτρους απόδοσης (χρόνος απόκρισης) πρέπει να κοστολογηθούν;
- Οι τιμές πρέπει να διαμορφώνονται ώστε η ζήτηση να ισούται με την προσφορά.
- Οι τιμές πρέπει να αντικατοπτρίζουν τη ζήτηση των καταναλωτών.
- Οι τιμές δεν πρέπει να μεταβάλλονται συνέχεια (υπάρχει κίνδυνος οι καταναλωτές να επιλέξουν λάθος προμηθευτές).

- Η διαπραγμάτευση των τιμών πρέπει να γίνεται πριν ληφθεί η απόφαση ”αγοράς” ενός αγαθού.

1.2 Οργάνωση εργασίας

Στην παρούσα εργασία έχουν σχεδιαστεί και μελετηθεί αλγόριθμοι κατανομής φόρτου και διαχείρισης δεδομένων που στηρίζονται στην οικονομία τιμών. Καταναλωτές της οικονομίας θεωρούνται δοσοληψίες, οι οποίες αγοράζουν αγαθά (υπολογιστικό χρόνο, εύρος επικοινωνίας) για να καλύψουν τις ανάγκες τους. Προμηθευτές της οικονομίας θεωρούνται οι κόμβοι του συστήματος, οι οποίοι εμπορεύονται τα τοπικά τους αγαθά και στοχεύουν στη μεγιστοποίηση των κερδών τους.

Συγκεκριμένα στην οικονομία κατανομής φόρτου εργασιών, οι καταναλωτές ομαδοποιούνται με βάση τις υπολογιστικές τους προτιμήσεις και αγοράζουν αγαθά, επιδιώκοντας την ελαχιστοποίηση του κόστους της ομάδας. Οι προμηθευτές διαμορφώνουν τις τιμές των αγαθών τους ανάλογα με την ζήτηση. Η ζήτηση εκφράζει τον φόρτο εργασιών που συσσωρεύεται σε κάθε κόμβο.

Στην οικονομία διαχείρισης δεδομένων οι καταναλωτές ομαδοποιούνται και αγοράζουν αγαθά επιδιώκοντας την ελαχιστοποίηση του κόστους της ομάδας. Ως κόστος θεωρείται πάλι ο μέσος χρόνος απόκρισης της ομάδας των δοσοληψιών. Οι προμηθευτές πουλούν τα τοπικά τους αγαθά, επιδιώκοντας την μεγιστοποίηση των προσωπικών τους κερδών. Οι προμηθευτές έχουν τη δυνατότητα σε τακτά χρονικά διαστήματα να ενοικιάζουν αγαθά (δεδομένα) τα οποία δεν διαθέτουν τοπικά και τα οποία πιστεύουν ότι θα τους αποφέρουν κέρδη στο άμεσο μέλλον. Η τιμή του ενοικίου είναι συνάρτηση της ζήτησης που παρατηρείται στο σύστημα για το συγκεκριμένο αγαθό. Οι τιμές των αγαθών και των αντιγράφων διαμορφώνονται με βάση τη ζήτηση των αγαθών (συσσωρευμένος φόρτος συστήματος) και την προσφορά (υπολογιστικές ικανότητες) των προμηθευτών.

Οι οικονομικοί αλγόριθμοι που σχεδιάστηκαν αποτέλεσαν τους αλγορίθμους δρομολόγησης σε πειράματα προσομοιώσεων, τα οποία διεξήχθηκαν στον προσομοιωτή **TPSim** (transaction processing simulator) [22]. Στα πειράματα αυτά χρησιμοποιήθηκαν διάφορα συστήματα επεξεργασίας δοσοληψιών. Η ανάλυση των αποτελεσμάτων των πειραμάτων οδήγησε στη διαμόρφωση συμπερασμάτων για την απόδοση και τη

συμπεριφορά (κάτω από διαφορετικές συνθήκες) των οικονομικών αλγορίθμων.

Η εργασία αυτή δομείται στα ακόλουθα κεφάλαια ως εξής: Στο Κεφάλαιο 2 περιγράφεται ο προσομοιωτής συστημάτων επεξεργασίας δοσοληψιών **TPSim** [22] και αναφέρονται οι προσθήκες που έγιναν στον προσομοιωτή ώστε να μπορεί να υποστηρίξει το οικονομικό περιβάλλον εργασίας.

Το Κεφάλαιο 3 αναφέρεται στην οικονομία κατανομής φόρτου. Συγκεκριμένα στο υποκεφάλαιο 3.1 περιγράφονται οι τρεις οικονομικοί αλγόριθμοι δρομολόγησης δοσοληψιών, ECO, COMM, SQRD. Στο υποκεφάλαιο 3.2 περιγράφονται τα συστήματα επεξεργασίας δοσοληψιών που χρησιμοποιήθηκαν στις προσομοιώσεις ενώ στο υποκεφάλαιο 3.3 αναλύονται τα αποτελέσματα τον προσομοιώσεων αυτών. Στο υποκεφάλαιο 3.4 χρησιμοποιείται ανάλυση στατικών μοντέλων με βάση τη θεωρία ουρών, για να πιστοποιηθούν κάποια από τα αποτελέσματα, ενώ το υποκεφάλαιο 3.5 αποτελεί επισκόπηση της σχετικής βιβλιογραφίας.

Το Κεφάλαιο 4 αναφέρεται στην οικονομία διαχείρισης δεδομένων. Η δόμησή του είναι ανάλογη μ'αυτή του Κεφαλαίου 3. Στο υποκεφάλαιο 4.1 περιγράφεται ο οικονομικός αλγόριθμος διαχείρισης δεδομένων και δρομολόγησης δοσοληψιών DRR, στο υποκεφάλαιο 4.2 περιγράφονται τα συστήματα επεξεργασίας δοσοληψιών που χρησιμοποιήθηκαν στις προσομοιώσεις, ενώ στο υποκεφάλαιο 4.3 αναλύονται τα αποτελέσματα των προσομοιώσεων αυτών. Στο υποκεφάλαιο 4.4 αποτελεί επισκόπηση της σχετικής βιβλιογραφίας.

Στο Κεφάλαιο 5 συνοψίζονται συμπεράσματα που αποκομίθηκαν από αυτή την εργασία και δίνονται οι επεκτάσεις και οι μελλοντικές ερευνητικές κατευθύνσεις αυτής της εργασίας.

Οι προσθήκες που πραγματοποιήθηκαν στον προσομοιωτή **TPSim** για την υποστήριξη ενός οικονομικού περιβάλλοντος εργασίας, ο σχεδιασμός και η μελέτη της οικονομίας κατανομής φόρτου και της οικονομίας διαχείρισης δεδομένων αποτελούν την συνεισφορά της παρούσας εργασίας.

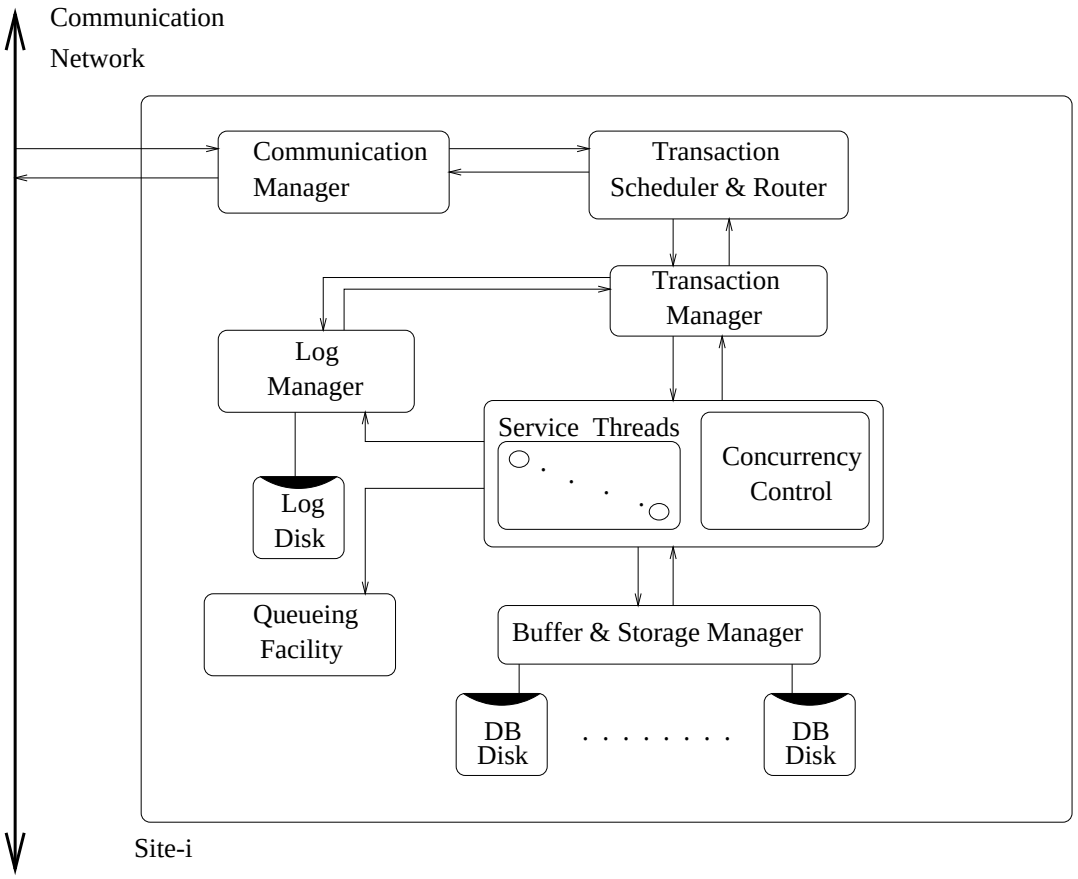
Κεφάλαιο 2

Ο Προσομοιωτής TPsim

Στο κεφάλαιο αυτό παρουσιάζεται ο προσομοιωτής κατανεμημένων συστημάτων επεξεργασίας δοσοληψιών **TPSim** (transaction processing simulator). Ο προσομοιωτής περιλαμβάνει μοντέλα για τους υπολογιστικούς κόμβους, τις συσκευές αποθήκευσης δεδομένων, και το δίκτυο επικοινωνίας, καθώς και για τα βασικά υποσυστήματα λογισμικού που απαρτίζουν ένα σύστημα επεξεργασίας δοσοληψιών, συμπεριλαμβανομένων και των διαχειριστών πόρων του συστήματος. Κεντρικό ρόλο στην αρχιτεκτονική του προσομοιωτή κατέχει ένας μεταφραστής για μια γλώσσα προδιαγραφής, με την οποία είναι η δυνατή η περιγραφή της διαμόρφωσης του συστήματος προς προσομοίωση και του προσφερόμενου φόρτου εξυπηρέτησης στο σύστημα. Ο μεταφραστής αυτός συνθέτει τελικά το μοντέλο του προσομοιούμενου συστήματος κάνοντας χρήση των μηχανισμών που παρέχει μια βιβλιοθήκη για την υποστήριξη προσομοίωσης. Ο προσομοιωτής εξομοιώνει συμβάντα που ανακύπτουν κατά την λειτουργία του υπό μελέτη συστήματος, και παράγει μία εκτίμηση για τις τιμές διαφόρων μεταβλητών του μοντέλου προσομοίωσης συστήματος και μιας σειράς από μέτρα επίδοσης. Ο προσομοιωτής εντάσσεται σε ένα ολοκληρωμένο περιβάλλον υποστήριξης πειραμάτων για την αυτοματοποίηση της διαδικασίας διεξαγωγής πειραμάτων προσομοίωσης και συλλογής μετρήσεων.

2.1 Το Μοντέλο Συστήματος

Ο προσομοιωτής Tpsim υλοποιεί μοντέλα για συστήματα επεξεργασίας δοσοληψιών. Ο προσομοιωτής περιλαμβάνει υποσυστήματα που μοντελοποιούν το υλικό (hardware) και το



Σχήμα 2.1: Το Μοντέλο Οργάνωσης ενός κόμβου

λογισμικό (software) ενός συστήματος επεξεργασίας δοσοληψιών. Το μοντέλο συστήματος περιλαμβάνει γενικά έναν επόπτη επεξεργασίας δοσοληψιών (transaction processing monitor) και έναν αριθμό από διαχειριστές πόρων (resource managers). Οι διαχειριστές πόρων που μοντελοποιούνται σε κάθε κόμβο του προσομοιουμένου συστήματος είναι ένα σύστημα διαχείρισης δεδομένων (database management system), ένα υποσύστημα επικοινωνίας (communication manager) και ένα σύστημα διαχείρισης ουρών (queueing system). Το μοντέλο συστήματος αναλύεται σε ένα σύνολο από υποσυστήματα. Το Σχήμα 2.1 απεικονίζει το μοντέλο οργάνωσης ενός κόμβου επεξεργασίας. Για κατανεμημένα συστήματα επεξεργασίας δοσοληψιών ακολουθείται η αρχιτεκτονική Shared-Nothing. Τα δεδομένα που προσπελαύνονται από τις δοσοληψίες που υποβάλλονται στο σύστημα είναι διαμερισμένα (partitioned) μεταξύ των κόμβων βάση ενός σχήματος ανάθεσης. Μια αίτηση προσπέλασης στη βάση δεδομένων από μια δοσοληψία που εκτελείται σε ένα

κόμβο, εξυπηρετείται τοπικά, εάν το ζητούμενο αντικείμενο δεδομένων έχει ανατεθεί στον κόμβο, ή μεταβιβάζεται για εξυπηρέτηση σε άλλο κόμβο που διαθέτει το ζητούμενο αντικείμενο δεδομένων. Η μεταβίβαση αίτησης προσπέλασης (function request shipping) είναι ο βασικός μηχανισμός για την κατανεμημένη εκτέλεση δοσοληψιών αφού επιτρέπει σε μία δοσοληψία που εκτελείται τοπικά να προσπελάσει πόρους ενός άλλου κόμβου.

2.2 Γλώσσα Προδιαγραφής

Η γλώσσα προδιαγραφής έχει ως στόχο να περιγράψει την διαμόρφωση του συστήματος καθώς και τον φόρτο εξυπηρέτησης που θα προσομοιωθεί ώστε το περιβάλλον εκτέλεσης που παρέχει ο προσομοιωτής να προσαρμοστεί στις προδιαγραφές της συγκεκριμένης μελέτης προσομοίωσης. Ο μεταφραστής αυτής της γλώσσας επεξεργάζεται αρχείο εισόδου που περιγράφει την διαμόρφωση του υπό μελέτη συστήματος, το σχήμα της κατανεμημένης βάσης δεδομένων, την ανάθεση τμημάτων της βάσης σε περιφερειακή μνήμη κόμβων του κατανεμημένου συστήματος, ποικιλία παραμέτρων κόστους για το σύστημα και επιπλέον μια περιγραφή του προσφερόμενου στο σύστημα φόρτου, μέσω του καθορισμού των τύπων δοσοληψιών που μπορούν να εμφανισθούν και της διατύπωσης απλών στατιστικών προδιαγραφής για την τοπικότητα των αναφορών σε δεδομένα. Η πληροφορία που παρέχεται από τον χρήστη του προσομοιωτή μέσω αυτού του αρχείου χρησιμοποιείται για την υλοποίηση των δομικών στοιχείων του μοντέλου. Παρακάτω περιγράφονται οι κύριες συντακτικές δομές της γλώσσας και δίνονται παραδείγματα χρήσης τους. Στα παραδείγματα με κεφαλαία σημειώνονται τα τελικά σύμβολα (tokens) της γλώσσας.

2.2.1 Περιγραφή Συσκευών Αποθήκευσης

Η γλώσσα προδιαγραφής παρέχει τη συντακτική δομή DEFINE_IO_DEVICE για τον καθορισμό παραμέτρων που επηρεάζουν την επίδοση των συσκευών αποθήκευσης. Το μοντέλο του συστήματος που υποστηρίζεται καθορίζει ότι κάθε κόμβος έχει ακριβώς μια συσκευή αποθήκευσης για καταγραφή αλλαγών ενώ μπορεί να έχει οποιοδήποτε πλήθος συσκευών αποθήκευσης συσκευών δεδομένων. Η κύρια κατηγορία συσκευών αποθήκευσης είναι οι μαγνητικοί δίσκοι (datadisk). Ακολουθεί παράδειγμα περιγραφής ενός τύπου δίσκων:

DEFINE_IO_DEVICE DataDisk WITH {

```
 IO_COPY_DELAY : 0.001;              %time to transfer a block to/from memory
 IO_LOAD_DELAY : 0.00075;            %time to "load" disk arm
 IO_SEEK_DELAY : 0.00075;            %time to move disk head from track to track
 IO_ROTATIONAL_DELAY : 0.00075;      %time for full disk rotation
 NUM_HEADS : 8;                      %number of disk heads
 NUM_SECTORS : 250000;               %number of blocks (total: around 2 GBytes)
}
```

Το χρόνο περίπτωσης μιας προσπέλασης καθορίζουν τρεις παράγοντες: η καθυστέρηση αναζήτησης (seek delay), η καθυστέρηση λόγω περιστροφής (rotational delay) και η καθυστέρηση μεταφοράς δεδομένων από και προς την κύρια μνήμη (I/O transfer delay). Η γλώσσα υποστηρίζει και μια απλουστευμένη μορφή περιγραφής ενός τύπου συσκευών που απαιτεί μόνο του καθορισμό (αναμενόμενης) ελάχιστης και μέγιστης καθυστέρησης προσπέλασης στην συσκευή. Κατά την προσομοίωση του συστήματος η καθυστέρηση για μια προσπέλαση λαμβάνεται να είναι μια τυχαία μεταβλητή με ομοιόμορφη κατανομή στο διάστημα που ορίζεται από τις ακραίες τιμές καθυστέρησης. Ακολουθεί παράδειγμα τέτοιας περιγραφής:

```
DEFINE_IO_DEVICE LogDisk WITH {
 IO_DELAY_MIN : 0.15;      %minimum I/O delay : 15 msec
 IO_DELAY_MAX : 0.020;     %maximum I/O delay : 0.020 msec
}
```

2.2.2 Περιγραφή Κόμβων και Δικτύου Κόμβων

Για την περιγραφή ενός κόμβου, η γλώσσα απαιτεί να έχει οριστεί ένας τύπος κόμβων, του οποίου ο κόμβος αποτελεί εμφάνιση (instance). Στα πλαίσια αυτής της εργασίας μελετήθηκαν συστήματα με μονοεπεξεργαστικούς (uniprocessor) κόμβους, αν και η γλώσσα προδιαγραφής επιτρέπει τον ορισμό κόμβων με περισσότερους επεξεργαστές οι οποίοι οργανώνονται σε μία αρχιτεκτονική τύπου UMA (Uniform Memory Access).

Τύποι Κόμβων

Η γλώσσα προδιαγραφής παρέχει τη συντακτική δομή DEFINE_NODE_CLASS για τη δήλωση ενός τύπου κόμβων. Ακολουθεί παράδειγμα τέτοιας δήλωσης:

```
DEFINE_NODE_CLASS nodeType WITH {
```

```
  CPUcnt : 1;          %number of CPUs
  MPL : 50;            %multiprogramming level
  CPUrate : 50.0;      %CPU capacity, measured in MIPS
%The following costs are expressed as instruction counts
  ATTACH_TASK_COST : 15100.0;              %cost of attaching a transaction with a thread
  DM_INTERFACE_COST : 2000.0;              %fixed cost for access to the DB
  DM_CALL_COST : 4000.0;                   %average cost for executing a DB access call
  DM_IO_COST : 10000.0;                    %fixed cost for access to an I/O device
  FUNCTION_SHIP_SEND_COST : 12600.0;       %cost of sending a remote request
  FUNCTION_SHIP_RECV_COST : 12600.0;       %cost of receiving a remote response
%The following four parameters define the cost of the z-phase commit protocol
%(coordinator side)
  PRIMARY_PREPARE_COST : 70000;
  SEND_PREPARE_COST : 12600.0;
  SEND_COMMIT_COST : 12600.0;
  PRIMARY_COMMIT_COST : 14000.0;
%The following four parameters define the cost for the z-phase commit protocol  %(participant
side)
  RECV_PREPARE_COST : 12600.0;
  RECV_COMMIT_COST : 12600.0;
  SECONDARY_PREPARE_COST : 12000.0;
  SECONDARY_COMMIT_COST : 12000.0;
  LOG_IO_COST : 5000.0;                    %cost of logging I/O
  DETACH_TASK_COST : 15100.0;              %cost of detaching a transaction from a thread
  DM_BUFFER_SIZE : 20000;                  %size of database buffer : 160 MBytes (8 KB pages)
  LOG_IO_DEVICE : LogDisk;                 %type of I/O device used for logging
  DM_IO_DEVICE : DataDisk;                 %type of I/O device(s) used for data sto rage
  numDisks : 2;                            %number of disks with data files
}
```

Ο κάθε επεξεργαστής χαρακτηρίζεται από το (μέσο) ρυθμό με τον οποίο εκτελεί εντολές, ο οποίος μετράται σε εκατομμύρια εντολών ανά second. Ο βαθμός πολυπρογραμματισμού καθορίζει το μέγιστο πλήθος δοσοληψιών που μπορούν να είναι ταυτόχρονα ενεργές σε έναν κόμβο. Ενα μέρος της μνήμης κάθε κόμβου δεσμεύεται για να χρησιμοποιηθεί ως

ενταμιευτής με σκοπό να μειώσει τις προσπελάσεις στις συσκευές αποθήκευσης. Κάθε κόμβος έχει μία συσκευή αποθήκευσης για την καταγραφή αλλαγών και ένα πλήθος από συσκευές αποθήκευσης δεδομένων. Η γλώσσα προδιαγραφής επιτρέπει τον καθορισμό του τύπου των συσκευών αποθήκευσης καθώς και τον αριθμό των συσκευών αποθήκευσης δεδομένων. Ολες οι συσκευές αποθήκευσης δεδομένων είναι πανομοιότυπες στα μοντέλα προσομοίωσης. Την περιγραφή ενός τύπου κόμβων συμπληρώνει ένας αριθμός από παραμέτρους που εκφράζουν το κόστος εκτέλεσης βασικών λειτουργιών του συστήματος επεξεργασίας δοσοληψιών.

Δήλωση Κόμβων

Εχοντας ορίσει ένα τύπο κόμβων, είναι δυνατή η δήλωση ενός συγκεκριμένου κόμβου, ως εμφάνιση του τύπου. Η δήλωση γίνεται ως εξής:

DEFINE NODE ProcessingNode OF CLASS nodeType;

Μια παραλλαγή αυτής της δήλωσης επιτρέπει τη δήλωση κόμβων που δεν εξυπηρετούν δοσοληψίες οι ίδιοι αλλά περιορίζονται στη δρομολόγηση δοσοληψιών. Η δήλωση ενός τέτοιου κόμβου (front_end) φαίνεται στο ακόλουθο παράδειγμα:

DEFINE NODE FrondEndNode OF CLASS nodeType (FROND_END);

Δήλωση Ομάδας Κόμβων

Η γλώσσα προδιαγραφής επιτρέπει τη δήλωση ομάδων κόμβων, όπου κάθε ομάδα κόμβων αποτελείται από ένα αριθμό μελών που διασυνδέονται μέσω ενός τοπικού δικτύου επικοινωνίας. Μία ομάδα κόμβων μπορεί να είναι ομογενής, να αποτελείται δηλαδή από ένα αριθμό κόμβων του ίδιου τύπου ή ετερογενής. Μια ετερογενής ομάδα κόμβων δηλώνεται με απαρίθμηση των μελών της. Η γλώσσα προδιαγραφής επιτρέπει τον καθορισμό του ρυθμού μετάδοσης και του μεγέθους πακέτων με τα οποία μεταδίδονται τα μηνύματα αλλά δεν παρέχει τη δυνατότητα καθορισμού κάποιας συγκεκριμένης τοπολογίας. Στα παρακάτω παραδείγματα φαίνεται η δήλωση μιας ομοιογενούς ομάδας κόμβων (με 8 μέλη) και μιας ετερογενούς ομάδας κόμβων (με 2 μέλη):
DEFINE NODE CLUSTER HomogeneousCluster OF CLASS nodeType WITH {

```
  numNodes : 8;            %number of nodes in cluster
  packetSize : 1024;       %packet size (in bytes)
  transferRate : 500000;   %measured in bytes per sec (4MBit)
}
DEFINE NODE NodeA OF CLASS nodeTypeA;
DEFINE NODE NodeB OF CLASS nodeTypeB;
    DEFINE NODE CLUSTER HeterogeneousCluster WITH {
  numNodes : 2;                    %number of nodes in cluster
  packetSize : 1024;               %packet size (in bytes)
  transferRate : 500000;           %measured in bytes per sec (4MBit)
  MEMBER_NODES : NodeA, NodeB
}
```

Δήλωση Κόστους Επικοινωνίας

Η γλώσσα προδιαγραφής παρέχει τη δυνατότητα να οριστούν σύνδεσμοι επικοινωνίας (communication links), για την απευθείας σύνδεση δύο κόμβων. Για κάθε μονόδρομο σύνδεσμο επικοινωνίας δύο κόμβων καθορίζεται η τιμή του. Η δήλωση του κόστους επικοινωνίας αποτελεί καινούργιο συντακτικό στοιχείο της γλώσσας προδιαγραφής.

```
DEFINE COMM_COST FROM node_1 TO node_2 : link_price;
```

2.2.3 Περιγραφή Βάσης Δεδομένων

Η γλώσσα προδιαγραφής επιτρέπει την περιγραφή της βάσης δεδομένων με τον ορισμό του σχήματος της βάσης δεδομένων (database schema) κατά το σχεσιακό μοντέλο, του σχήματος τμηματοποίησης (fragmentation schema) και του σχήματος ανάθεσης (allocation scheme). Το σχήμα της βάσης δεδομένων ορίζεται καθορίζοντας το σχήμα κάθε σχέσης. Παράδειγμα ορισμού σχήματος σχέσης δίδεται παρακάτω. Απαριθμούνται τα πεδία (attributes) κάθε σχέσης, καθορίζεται ο τύπος για το καθένα από αυτά, και ορίζεται το πρωτεύον κλειδί της σχέσης. Είναι ακόμα δυνατός ο προσδιορισμός πεδίων βάσει των οποίων το σύστημα διαχείρισης δεδομένων μπορεί να κατασκευάσει δομή δεικτοδότησης (index)

```
DEFINE RELATION R WITH {
    ATTRIBUTES
```

```
        rA : SYMBOLIC;
        rB : NUMERIC;
    tuplesPerPage : 20.0;
    KEY = rA;
    INDEX ON = rA;
}
```

Για κάθε σχέση επιτρέπεται ο καθορισμός τμηματοποίησης που μπορεί να είναι είτε οριζόντια, είτε κατακόρυφος. Με τον ορισμό τμηματοποίησης μια σχέση αντιστοιχεί σε ένα ή περισσότερα τμήματα. Για ένα τμήμα μιας σχέσης αλλά και για μια σχέση είναι δυνατόν να οριστούν πολλαπλά αντίγραφα (replicas). Κάθε τμήμα σχέσης και κάθε σχέση υλοποιείται ως ένα αρχείο δεδομένων και προσδιορίζονται οι κόμβοι στους οποίους αποθηκεύονται τα αντίστοιχα αρχεία δεδομένων. Το παρακάτω παράδειγμα ορίζει το σχήμα ανάθεσης για δύο σχέσεις R, S όπου για την R ορίζεται οριζόντια τμηματοποίηση και ένα αντίγραφο για κάθε τμήμα, ενώ για την S ορίζονται δύο αντίγραφα χωρίς να έχει οριστεί τμηματοποίηση. Καθορίζεται επίσης και η τιμή κάθε αντιγράφου. Η δυνατότητα ορισμού πολλαπλών αντιγράφων καθώς και η κοστολόγηση κάθε αντικειμένου δεδομένου αποτελούν καινούργια συντακτικά στοιχεία της γλώσσας προδιαγραφής.

```
DEFINE HORIZONTAL FRAGMENATION OF R AS {
    TblH1 WITH ((rB >= 128) OR (rB < 512)),
        tuplesPerPage : 20.0;
    TblH2 WITH ((rB >= 128) OR (rB >= 512)),
        tuplesPerPage : 20.0;
}
```

```
    Allocation of fragments of R
DEFINE INSTANCE OF TblH1 AT {Node1 WITH VALUE : value1}
DEFINE INSTANCE OF TblH2 AT {Node2 WITH VALUE : value2}
```

```
    Allocation of S
DEFINE INSTANCE OF S AT {Node1 WITH VALUE : value1, Node2 WITH VALUE : value2}
WITH numItems : 1000;
```

2.2.4 Περιγραφή Φόρτου Εξυπηρέτησης

Για την περιγραφή του φόρτου εξυπηρέτησης (workload), η γλώσσα προδιαγραφής παρέχει συντακτικές δομές για τον ορισμό κλάσεων μονάδων φόρτου εξυπηρέτησης (classes of units of work) και κλάσεων χρηστών. Η γλώσσα προδιαγραφής επιτρέπει τον ορισμό δύο τύπων κλάσεων μονάδων φόρτου. Οι απλές μονάδες φόρτου είναι δοσοληψίες που διαβάζουν και ενημερώνουν δεδομένα. Για την περιγραφή μιας απλής κλάσης μονάδων φόρτου καθορίζεται το (αναμενόμενο) πλήθος εντολών που εκτελεί το αντίστοιχο πρόγραμμα εφαρμογής, και η κατανομή των προσπελάσεων από δοσοληψίες αυτής της κλάσης στα αρχεία που υλοποιούν στο φυσικό επίπεδο τη βάση δεδομένων. Καθορίζεται το ελάχιστο και το μέγιστο πλήθος προσπελάσεων, καθώς και η πιθανότητα προσπέλασης για κάθε αρχείο δεδομένων που οι δοσοληψίες αυτής της κλάσης χρειάζονται. Για κάθε προσπελαζόμενο αρχείο δεδομένων καθορίζεται επίσης η πιθανότητα η προσπέλαση να είναι εγγραφή. Καθορίζεται τέλος το κεφάλαιο που διατίθεται σε κάθε δοσοληψία αυτής της κλάσης. Η δήλωση του κεφαλαίου αποτελεί καινούργιο συντακτικό στοιχείο της γλώσσας προδιαγραφής. Ακολουθεί παράδειγμα δήλωσης της απλής κλάσης classA1.

```
DEFINE TRANSACTION_CLASS classA1 AS {
    applicationBurst : 80000.0;
    blocksAccessed_min : 4;
    blocksAccessed_max : 12;
    budget : 500000;
    accessProbability OF INSTANCE rA : 0.2;
    writeProbability OF INSTANCE rA : 0.8;
    accessProbability OF INSTANCE rB : 0.2;
    writeProbability OF INSTANCE rA : 0.8;
}
```

Μια σύνθετη μονάδα φόρτου (workflow) ορίζεται ως σύνθεση απλών μονάδων φόρτου δοσοληψιών. Η γλώσσα προδιαγραφής υποστηρίζει τον ορισμό αλυσίδων από δοσοληψίες. Ακολουθεί παράδειγμα δήλωσης σύνθετης μονάδας φόρτου. Η κλάση δοσοληψιών classB θεωρείται μια αλυσίδα από 3 δοσοληψίες που ανήκουν στις κλάσεις classA1, classA2, classA3.

```
DEFINE WORKFLOW_CLASS WC1 AS {
    programm : classB_Program;
```

```
    chain : classA1, classA2, classA3
}
```

Το ακόλουθο παράδειγμα περιγράφει τον ορισμό μιας κλάσης χρηστών που υποβάλλουν στο σύστημα δοσοληψίες των κλάσεων classA1, classB με την ίδια πιθανότητα. Καθορίζεται επίσης ο ρυθμός άφιξης δοσοληψιών ο οποίος μπορεί να είναι σταθερός ή να μεταβάλλεται στη διάρκεια του χρόνου προσομοίωσης. Στο παράδειγμα ο ρυθμός άφιξης δοσοληψιών είναι σταθερός στο χρόνο.

```
DEFINE CLIENT_CLASS clientClass AS {
interarrivalRate : 5.0;
    SUBMIT classA1 WITH PROBABILITY : 0.5;
    SUBMIT classB WITH PROBABILITY : 0.5;
}
```

Για κάθε κλάση χρηστών καθορίζεται ο αριθμός των ενεργών χρηστών σε κάθε κόμβο του συστήματος.

```
DEFINE SYNTHETIC WORKLOAD wrkLoad AS {
    numClients OF CLASS clientClass : 50 AT {Node1, Node2}
}
```

Σε όλα τα πειράματα προσομοίωσης που διεξήχθησαν στα πλαίσια αυτής της εργασίας χρησιμοποιήθηκε το ανοιχτό μοντέλο φόρτου στο οποίο θεωρείται ότι υπάρχει, σε κάθε κόμβο του συστήματος και για κάθε κλάση χρηστών που είναι ενεργή στον κόμβο, μια πηγή-φόρτου που στέλνει αιτήσεις εξυπηρέτησης, με ρυθμό που καθορίζεται ως παράμετρος στον ορισμό της κλάσης χρηστών. Η πηγή φόρτου δεν περιμένει να λάβει ειδοποίηση για την περάτωση μιας δοσοληψίας για να αποστείλει την επόμενη.

2.3 Προσθήκες στο Tpsim για την υποστήριξη οικονομικών αλγορίθμων

Για την εισαγωγή της έννοιας της οικονομίας στο περιβάλλον προσομοίωσης ήταν απαραίτητος ο εμπλουτισμός της γλώσσας προδιαγραφής. Συγκεκριμένα, η γλώσσα επιτρέπει στα πλαίσια περιγραφής της βάσης δεδομένων τον καθορισμό της τιμής κάθε σχέσης, τμήματος σχέσης ή αντιγράφου στους κόμβους στους οποίους αποθηκεύονται τα αντίστοιχα αρχεία δεδομένων. Επίσης στα πλαίσια περιγραφής μιας απλής κλάσης

μονάδων φόρτου επιτρέπεται ο καθορισμός του κεφαλαίου που χορηγείται σε κάθε δοσοληψία της κλάσης. Η γλώσσα προδιαγραφής παρέχει επιπλέον τη δυνατότητα ορισμού της τιμής συνδέσμου επικοινωνίας για την απευθείας σύνδεση δύο κόμβων. Η τιμή κάθε σχέσης, τμήματος σχέσης ή αντιγράφου, το κεφάλαιο κάθε κλάσης δοσοληψιών καθώς και η τιμή κάθε συνδέσμου επικοινωνίας τείθονται από τον χρήστη του προσομοιωτή.

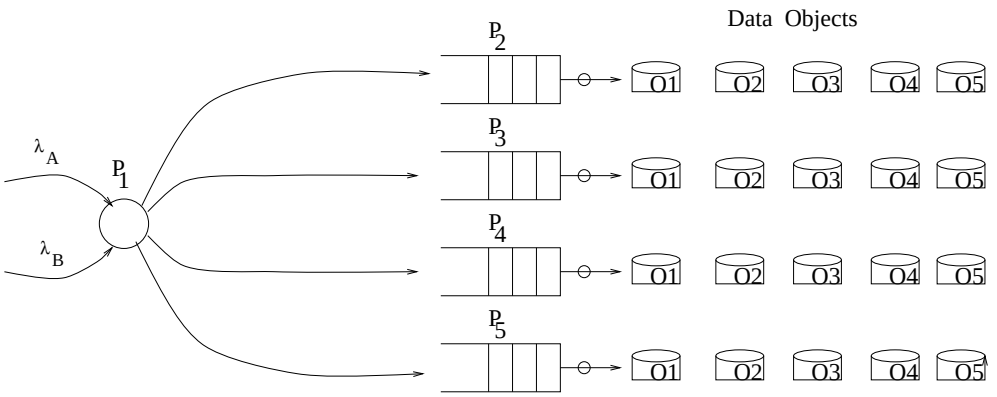
Εκτός από τον εμπλουτισμό της γλώσσας προδιαγραφής στον προσομοιωτή TPsim προστέθηκαν μηχανισμοί για την υποστήριξη πολλαπλών αντιγράφων μιας σχέσης ή ενός τμήματος σχέσης. Δίδεται πλέον η δυνατότητα δημιουργίας ενός αντιγράφου ανάγνωσης μιας σχέσης ή τμήματος σχέσης σε έναν κόμβο κατά την εξέλιξη της προσομοίωσης αλλά και η δυνατότητα απόρριψης ενός αντιγράφου από τις συσκευές αποθήκευσης δεδομένων του κόμβου στον οποίο έχει ανατεθεί.

Η ανταλλαγή πληροφοριών μεταξύ των κόμβων γίνεται μέσω καθολικών πινάκων. Υπάρχουν πίνακες στους οποίους οι κόμβοι διαφημίζουν ποιά αντικείμενα δεδομένων διαθέτουν και σε ποιές τιμές. Υπάρχουν επίσης κατάλογοι (καθολικοί πίνακες) στους οποίους οι κόμβοι καταγράφουν τη συνολική ζήτηση που έχουν καθώς και τη ζήτηση για κάθε αντικείμενο δεδομένων που τους έχει ανατεθεί. Με την βοήθεια αυτών των καταλόγων υπολογίζεται κάθε χρονική στιγμή η αξία του υπολογιστικού χρόνου των κόμβων του συστήματος. Υπάρχουν επίσης κατάλογοι στους οποίους καταγράφεται το κόστος επικοινωνίας των κόμβων του συστήματος. Οι κατάλογοι αυτοί ενημερώνονται από τους κόμβους του συστήματος μόλις σημειωθεί αλλαγή ώστε οι τιμές των προϊόντων που διαφημίζουν να είναι πάντα επίκαιρες.

Κεφάλαιο 3

Οικονομία Κατανομής Φόρτου Εργασιών

3.1 Περιγραφή Προβλήματος



Σχήμα 3.1: Παράδειγμα Μοντέλου Προβλήματος

Ας θεωρήσουμε το κατανεμημένο σύστημα (Σχήμα 3.1) που αποτελείται από 5 επεξεργαστές P_1, P_2, P_3, P_4, P_5. Η υπολογιστική δύναμη του επεξεργαστή P_i καθορίζεται από την παράμετρο r_i όπου $i \in [1,5]$. Ο επεξεργαστής P_1 περιορίζεται στη δρομολόγηση δοσοληψιών σε άλλους κόμβους ενώ οι υπόλοιποι επεξεργαστές εξυπηρετούν δοσοληψίες. Οι επεξεργαστές συνδέονται με ένα ‘‘point-to-point’’ δίκτυο το οποίο καθορίζεται από το σύνολο ακμών $E = \{e_{ij}\}$ όπου $i, j \in [1,5]$ και $i \neq j$. Η ακμή e_{ij} παριστά ένα μονόδρομο

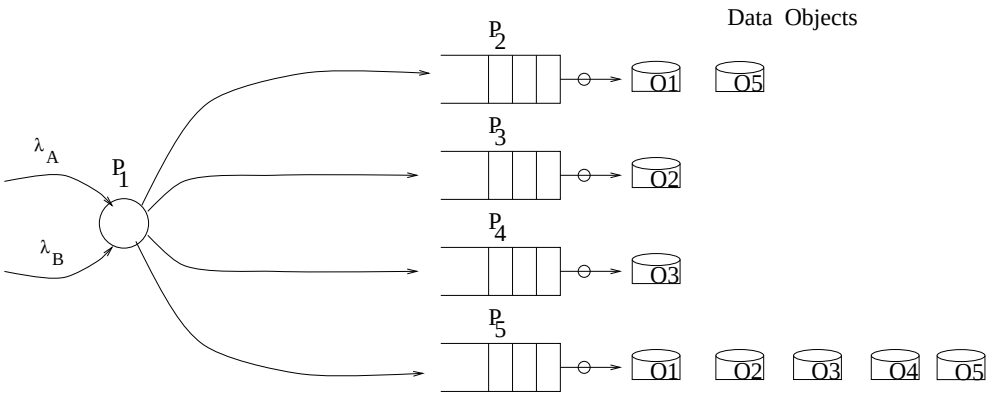
σύνδεσμο επικοινωνίας μεταξύ των επεξεργαστών P_i, P_j. Ο σύνδεσμος αυτός παρουσιάζει καθυστέρηση d_{ij} seconds per byte. Η βάση δεδομένων αποτελείται από 5 σχέσεις O_1, O_2, O_3, O_4, O_5 αντίγραφα των οποίων έχουν ανατεθεί σε όλους τους κόμβους. Εχουν οριστεί δύο κλάσεις δοσοληψιών C_A, C_B. Οι δοσοληψίες της κλάσης C_A προσπελαύνουν κάποιο από τα δεδομένα O_1, O_3, O_5 ενώ οι δοσοληψίες της κλάσης C_B προσπελαύνουν κάποιο από τα δεδομένα O_1, O_2, O_4, O_5. Θεωρούμε ότι οι υπολογιστικές απαιτήσεις των δοσοληψιών είναι γνωστές κατά την είσοδό τους στο σύστημα. Ο υπολογιστικός χρόνος που απαιτείται για την εξυπηρέτηση δοσοληψίας της κλάσης C_A σε ιδανικό επεξεργαστή P_i $(r_i = 1)$ είναι μ_A ενώ αντίστοιχα μ_B είναι ο χρόνος που απαιτείται για την εξυπηρέτηση δοσοληψίας της κλάσης C_B. Ο επεξεργαστής P_1 (πηγή φόρτου) παράγει δοσοληψίες των κλάσεων C_A, C_B με ρυθμούς λ_A, λ_B αντίστοιχα.

Το πρόβλημα της κατανομής του φόρτου των εργασιών (load balancing problem) σε αυτό το κατανεμημένο σύστημα ορίζεται ως η ελαχιστοποίηση του μέσου χρόνου απόκρισης (average responce time) των κλάσεων C_A, C_B. Ο χρόνος απόκρισης μιας δοσοληψίας j (της κλάσης C_j) σε ένα επεξεργαστή P_i ορίζεται ως το άθροισμα του χρόνου αναμονής της δοσοληψίας j (waiting time) και του χρόνου εξυπηρέτησης της δοσοληψίας (service time) στον επεξεργαστή P_i. Δεδομένα του προβλήματος θεωρούνται οι ρυθμοί άφιξης λ_A, λ_B των δύο κλάσεων καθώς και οι χρόνοι εξυπηρέτησης $\frac{\mu_A}{r_i}, \frac{\mu_B}{r_i}$ στον επεξεργαστή P_i δοσοληψιών των κλάσεων C_A, C_B αντίστοιχα.

Αναζητείται λοιπόν αλγόριθμος δρομολόγησης ο οποίος επιτυγχάνει επίλυση του υπό μελέτη προβλήματος. Ενας απλός παραδοσιακός αλγόριθμος δρομολόγησης που θα μπορούσε να χρησιμοποιηθεί είναι η ισοκατανομή του πλήθους των δοσοληψιών στους διαθέσιμους επεξεργαστές. Η ανάθεση των δεδομένων στο υπό μελέτη σύστημα (πλήρη αντιγραφή των σχέσεων της βάσης σε όλους τους κόμβους) επιτρέπει την εφαρμογή του αλγορίθμου αυτού. Πιο πολύπλοκες και δικαιότερες προσεγγίσεις λαμβάνουν υπόψη την υπολογιστική δύναμη και τον υπάρχοντα φόρτο των επεξεργαστών, τον αναμενόμενο χρόνο εξυπηρέτησης της δρομολογούμενης δοσοληψίας (service time) σε κάθε επεξεργαστή καθώς και την καθυστέρηση για τη μεταβίβαση της εργασίας στους επεξεργαστές. Οι παραπάνω παράμετροι είναι απαραίτητες για την ισοκατανομή του φόρτου του συστήματος όταν οι κλάσεις δοσοληψιών που έχουν οριστεί προσπελαύνουν διαφορετικά δεδομένα, έχουν διαφορετικό ρυθμό άφιξης και διαφορετικό χρόνο εξυπηρέτησης και οι κόμβοι του συστήματος έχουν διαφορετική υπολογιστική δύναμη, αποθηκευτική

χωρητικότητα, κ.λ.π..

Αν οι σχέσεις της βάσης δεδομένων δεν είχαν αντιγραφεί σε όλους τους κόμβους του συστήματος το πρόβλημα της ελαχιστοποίησης του μέσου χρόνου απόκρισης των κλάσεων δοσοληψιών C_A, C_B θα ήταν πιο πολύπλοκο. Σε ανάλογη περίπτωση ο αλγόριθμος δρομολόγησης θα έπρεπε να γνωρίζει ποιό δεδομένο επιθυμεί να προσπελάσει κάθε δοσοληψία καθώς και τους κόμβους στους οποίους έχει ανατεθεί το δεδομένο αυτό.



Σχήμα 3.2: Παράδειγμα Μοντέλου Προβλήματος

Ας θεωρήσουμε ότι στο προηγούμενο σύστημα οι σχέσεις της βάσης δεδομένων δεν έχουν αντιγραφεί σε όλους τους κόμβους του συστήματος. Συγκεκριμένα, οι σχέσεις O_1, O_5 έχουν ανατεθεί στον κόμβο P_2, οι σχέσεις O_2, O_3 στους κόμβους P_3, P_4 αντίστοιχα ενώ αντίγραφα όλων των σχέσεων της βάσης έχουν αποθηκευτεί στον κόμβο P_5 (Σχήμα 3.2). Για λόγους απλοποίησης του προβλήματος κάνουμε επίσης τις ακόλουθες παραδοχές

- οι ρυθμοί άφιξης των δοσοληψίων των κλάσεων C_A, C_B είναι ίσοι (λ_A, λ_B)
- οι δοσοληψίες των κλάσεων C_A, C_B απαιτούν σε ιδανικό επεξεργαστή τον ίδιο χρόνο εξυπηρέτησης (μ_A, μ_B)
- οι επεξεργαστές διαθέτουν την ίδια υπολογιστική ισχύ $r_i = r_j \forall i, j \in [1, 5]$

Το πρόβλημα της ελαχιστοποίησης του μέσου χρόνου απόκρισης των κλάσεων δοσοληψιών C_A, C_B δεν επιλύεται με ισοκατανομή των εργασιών στους 4 επεξεργαστές P_2, P_3, P_4, P_5. Οι επεξεργαστές P_2, P_5 που διαθέτουν περισσότερα δεδομένα έχουν τη δυνατότητα να εξυπηρετήσουν περισσότερες δοσοληψίες και να επιτύχουν (κατά κανόνα)

μικρότερο χρόνο απόκρισης για τις δοσοληψίες αυτές (για κάθε επεξεργαστή ορίζεται ο βαθμός πολυπρογραμματισμού που καθορίζει το μέγιστο πλήθος δοσοληψιών που μπορούν να είναι ταυτόχρονα ενεργές στον επεξεργαστή).

Ο αλγόριθμος δρομολόγησης που θα μπορούσε να επιλύσει το πρόβλημα θα έπρεπε να γνωρίζει το φόρτο που έχει συσσωρευτεί σε κάθε κόμβο, για κάθε μία από τις σχέσεις που έχουν ανατεθεί στον κόμβο (ουρές αναμονής για κάθε σχέση του κόμβου).

Ακόμη κι αν σχεδιαστεί αλγόριθμος δρομολόγησης που λαμβάνει υπόψη τις αναφερόμενες παραμέτρους και επιλύει το πρόβλημα της κατανομής του φόρτου εργασιών σε ένα κατανεμημένο σύστημα ένα διαφορετικό σχήμα ανάθεσης των δεδομένων μπορεί να οδηγήσει σε βελτιωμένη λύση (μικρότερο μέσο χρόνο απόκρισης των κλάσεων δοσοληψιών).

Στο κεφάλαιο αυτό ασχολούμαστε με τη σχεδίαση οικονομικών αλγορίθμων δρομολόγησης οι οποίοι εφαρμόζονται σε κατανεμημένο σύστημα για το οποίο καθορίζονται διαφορετικά σχήματα ανάθεσης. Η μελέτη των αποτελεσμάτων των προσομοιώσεων που διεξάγονται μας δίνουν τη δυνατότητα να εντοπίσουμε το συνδυασμό αλγορίθμου δρομολόγησης και σχήματος ανάθεσης για τον οποίο επιτυγχάνεται ελάχιστος μέσος χρόνος απόκρισης των κλάσεων δοσοληψιών του συστήματος, και να μελετήσουμε τη συμπεριφορά των οικονομικών αλγορίθμων.

3.2 Περιγραφή Οικονομικών Αλγορίθμων Δρομολόγησης

Οι οικονομικοί αλγόριθμοι αντιμετωπίζουν το κατανεμημένο σύστημα ως ανθρώπινη οικονομία. Οι κόμβοι του υπολογιστικού συστήματος διαθέτουν αγαθά (υπολογιστική δύναμη, εύρος ζώνης επικοινωνίας) τα οποία εμπορεύονται με στόχο τη μεγιστοποίηση των προσωπικών τους κερδών. Οι κλάσεις δοσοληψιών διαθέτουν ένα χρηματικό ποσό (κεφάλαιο) που το ισοκατανέμουν στις δοσοληψίες τους. Στόχος των κλάσεων δοσοληψιών είναι η εξυπηρέτηση των εργασιών τους με το μικρότερο εφικτό κόστος. Το κεφάλαιο κάθε κλάσης καθορίζεται από το διαχειριστή του συστήματος και αντικατοπτρίζει την προτεραιότητα και τις υπολογιστικές απαιτήσεις των δοσοληψιών της κλάσης. Οι κλάσεις δοσοληψιών έχουν τις ίδιες προτιμήσεις σε αγαθά (utility function) αλλά μπορεί να διαθέτουν διαφορετικά κεφάλαια. Οταν μια δοσοληψία εισέρχεται στο σύστημα χρησιμοποιεί το κεφάλαιο της για την αγορά υπολογιστικού χρόνου (CPU time) σε

κάποιον από τους κόμβους στον οποίο έχει ανατεθεί το δεδομένο που ζητά να προσπελάσει η δοσοληψία. Η εργασία δρομολογείται στον κόμβο που προσφέρει μικρότερη τιμή για την εξυπηρέτησή της. Η τιμή αποτελεί συνάρτηση του του φόρτου του συστήματος. Η δοσοληψία αποτυγχάνει να εξυπηρετηθεί όταν δε διαθέτει ισχυρό κεφάλαιο (όλοι οι κόμβοι ζητούν περισσότερα από όσα δύναται να πληρώσει). Το κεφάλαιο που διαθέτει κάθε κλάση αποτελεί ένα δείκτη προτεραιότητας. Οσο μεγαλύτερο είναι το κεφάλαιο μιας κλάσης τόσο μεγαλύτερο ποσοστό ενός αγαθού μπορεί να αγοράσει η κλάση αυτή και τόσο μικρότερη είναι η πιθανότητα απόρριψης δοσοληψιών της κλάσης αυτής (abort). Η οικονομική πολιτική που ακολουθείται για την διαμόρφωση των τιμών είναι κοινή για όλους τους κόμβους του συστήματος. Κανένας κόμβος δεν υιοθετεί δόλια συμπεριφορά, δε χρεώνει δηλαδή τα αγαθά του σε τιμές δυσανάλογες της αξίας τους. Ακολουθεί περιγραφή των οικονομικών αλγορίθμων δρομολόγησης για κατανομή του φόρτου εργασιών ενός συστήματος.

Γενίκευση Συστήματος

Θεωρώ σύστημα Ν επεξεργαστών $\{P_1, ..., P_N\}$ όπου r_i υπολογιστική δύναμη (CPU rate) του επεξεργαστή P_i για $i \in [1, N]$. Η βάση δεδομένων αποτελείται από Μ σχέσεις δεδομένων $\{\mathrm{O}_1, ..., \mathrm{O}_M\}$ και $S(\mathrm{O}_i)$ το πλήθος των σελίδων του δεδομένου O_i, $i \in [1, M]$. Με e_{ij} συμβολίζεται ο σύνδεσμος επικοινωνίας (communication link) μεταξύ των επεξεργαστών P_i, P_j και d_{ij} είναι η καθυστέρηση του συνδέσμου επικοινωνίας e_{ij} . Εχουν οριστεί Λ κλάσεις δεδομένων ($C_\mathrm{A}, C_\mathrm{B}, ..., C_\Lambda$) και μ_i ο χρόνος εξυπηρέτησης (service time) δοσοληψίας της κλάσης C_i σε ιδανικό επεξεργαστή.

- **Αλγόριθμος ECO**

  Οταν μια δοσοληψία j (της κλάσης C_j) εισέρχεται στο σύστημα επιθυμεί να αγοράσει υπολογιστικό χρόνο. Οι κόμβοι του συστήματος πωλούν υπολογιστικό χρόνο στη δοσοληψία η οποία θα επιλέξει να εξυπηρετηθεί στον κόμβο που ζητά τη μικρότερη τιμή για τις υπηρεσίες του. Ο κόμβος πρέπει να διαθέτει το δεδομένο το οποίο επιθυμεί να προσπελάσει η δοσοληψία. Η συνάρτηση χρησιμότητας αντανακλά τις προτιμήσεις της κλάσης C_j και είναι ίδια για όλες τις κλάσεις δοσοληψιών που έχουν οριστεί.

  Utility Function =

$$\min_{\forall k}\{C_k\}, k \in [1, N]$$

C_k είναι η τιμή που προσφέρει ο επεξεργαστής P_k

$$\min_{\forall k}\{C_k\} = \min_{\forall k}\{J * (ST_k + WT_k)\}$$

Η τιμή που προσφέρει ο κόμβος P_k είναι μια συνάρτηση του χρόνου αναμονής WT_k και του αναμενόμενου χρόνου εξυπηρέτησης ST_k της δοσοληψίας j στον κόμβο. Ο χρόνος αναμονής ορίζεται ως το άθροισμα των αναμενομένων χρόνων εξυπηρέτησης των δοσοληψιών που έχουν δρομολογηθεί στον επεξεργαστή, βρίσκονται στην ουρά και περιμένουν να εξυπηρετηθούν. Ο πολλαπλασιαστής J έχει τιμή 1 δραχμή ανά δευτερόλεπτο.

Ο αλγόριθμος λοιπόν δρομολογεί τη δοσοληψία στον επεξεργαστή που προσφέρει μικρότερο κόστος εξυπηρέτησης (αναμενόμενος χρόνος απόκρισης). Αν μια δοσοληψία δε διαθέτει χρηματικό ποσό ανάλογο των απαιτήσεων κάποιου κόμβου απορρίπτεται από το σύστημα (abort).

- **Αλγόριθμος COMM**

  Οταν μια δοσοληψία της j (κλάσης C_j) εισέρχεται στο σύστημα επιθυμεί να αγοράσει υπολογιστικό χρόνο. Οι επεξεργαστές του συστήματος πωλούν υπολογιστικό χρόνο. Η δοσοληψία θα επιλέξει να εξυπηρετηθεί στον κόμβο που διαθέτει το δεδομένο που θέλει να προσπελάσει και ζητά τη μικρότερη τιμή για τις υπηρεσίες του. Οι κλάσεις δοσοληψιών έχουν την ίδια συνάρτηση χρησιμότητας που αντανακλά τις προτιμήσεις τους.

  Utility Function =

$$\min_{\forall k}\{C_k\} = \min_{\forall k}\{J * (ST_k + WT_k + d_{ak})\} = \min_{\forall k}\{J * (\frac{\mu_j}{r_k} + WT_k + d_{ak})\}, k \in [1, N]$$

  C_k είναι η τιμή που προσφέρει ο επεξεργαστής P_k. Η τιμή αυτή είναι συνάρτηση του χρόνου αναμονής και εξυπηρέτησης της δοσοληψίας στον επεξεργαστή P_k καθώς και του χρόνου για τη μεταβίβαση της δοσοληψίας από την πηγή φόρτου P_a στον επεξεργαστή P_k. Ο πολλαπλασιαστής J έχει τιμή 1 δραχμή ανά δευτερόλεπτο.

  Ο αλγόριθμος δρομολογεί την εργασία στον επεξεργαστή που διαθέτει το δεδομένο προς προσπέλαση και προσφέρει το ελάχιστο άθροισμα κόστους εξυπηρέτησης (αναμενόμενος χρόνος απόκρισης) και κόστους μεταβίβασης της δοσοληψίας. Αν

μια δοσοληψία δε διαθέτει αρκετό κεφάλαιο ώστε να μπορεί να αγοράσει αγαθά από κάποιο κόμβο, απορρίπτεται από το σύστημα (abort).

- **Αλγόριθμος SQRD**

  Όταν μια δοσοληψία j (της κλάσης C_j) εισέρχεται στο σύστημα επιθυμεί να αγοράσει υπολογιστικό χρόνο σε επεξεργαστή που διαθέτει το δεδομένο που επιθυμεί να προσπελάσει. Η δοσοληψία θα δρομολογηθεί στον επεξεργαστή που ζητά τη μικρότερη αμοιβή για τις υπηρεσίες του. Η συνάρτηση χρησιμότητας που αντανακλά τις προτιμήσεις των δοσοληψιών είναι ίδια για όλες τις κλάσεις που έχουν οριστεί.

  Utility Function =

  $$\min_{\forall k}\{C_k\} = \min_{\forall k}\{K * (ST_k + WT_k)^2 + J * d_{aj}\} = \min_{\forall k}\{K * (\frac{\mu_j}{r_k} + WT_k)^2 + J * d_{aj}\}$$

  C_k είναι η τιμή που προσφέρει ο επεξεργαστής P_k. Η τιμή αυτή είναι συνάρτηση του χρόνου αναμονής WT_k και του αναμενόμενου χρόνου εξυπηρέτησης ST_k της δοσοληψίας j καθώς και του χρόνου για τη μεταβίβαση της δοσοληψίας από την πηγή φόρτου P_a στον επεξεργαστή P_k. Ο πολλαπλασιαστής J έχει τιμή 1 δραχμή ανά δευτερόλεπτο ενώ ο πολλαπλασιαστής K έχει τιμή 1 δραχμή ανά (δευτερόλεπτο)^2

  Ο αλγόριθμος δρομολογεί την εργασία στον επεξεργαστή που διαθέτει το δεδομένο προς προσπέλαση και προσφέρει το ελάχιστο άθροισμα του τετραγώνου του κόστους εξυπηρέτησης (αναμενόμενος χρόνος απόκρισης) της δοσοληψίας και του κόστους για τη μεταβίβασή της. Η διαφορά του αλγορίθμου αυτού από αλγόριθμο COMM είναι ότι στην κοστολόγηση των υπηρεσιών του ο επεξεργαστής δίνει μεγαλύτερη βαρύτητα στον αναμενόμενο χρόνο απόκρισης της δοσοληψίας. Το κόστος μεταβίβασης επηρεάζει λιγότερο την επιλογή του ‘‘πιο φθηνού’’ επεξεργαστή.

Στους τρεις αλγορίθμους που περιγράφηκαν οι τιμές των αγαθών κάθε κόμβου αποτελούν συνάρτηση της ζήτησης που παρατηρείται στον κόμβο αυτό. Σε κάθε αλγόριθμο ο χρόνος αναμονής της δοσοληψίας j στον επεξεργαστή P_k αποτελεί παράμετρο για την κοστολόγηση των υπηρεσιών του επεξεργαστή. Η παράμετρος αυτή αντικατοπτρίζει τη ζήτηση αφού αποτελεί μέτρο του υπολογιστικού φόρτου που έχει ήδη δρομολογηθεί στον επεξεργαστή. Οι τιμές των αγαθών ενός επεξεργαστή διαμορφώνονται ώστε να αντικατοπτρίζουν και τις δυνατότητες προσφοράς του επεξεργαστή. Στην κοστολόγηση των υπηρεσιών του ένας επεξεργαστής P_k λαμβάνει υπόψη του τόσο το χρόνο αναμονής μιας καινούργιας δοσοληψίας όσο και το χρόνο εξυπηρέτησής της.

Παράμετρος για τον υπολογισμό των χρόνων αυτών αποτελεί η υπολογιστική δύναμη (CPU rate) του επεξεργαστή. Όσο μικρότερη τιμή λοιπόν προσφέρει ο επεξεργαστής P_k τόσο μικρότερο υπολογιστικό φόρτο διαθέτει (σε σχέση με την υπολογιστική του δύναμη) και τόσο περισσότερο υπολογιστικό χρόνο (CPU time) δύναται να διαθέσει σε μια καινούργια δοσοληψία. Οι επεξεργαστές εγγυώνται ότι δεν πρόκειται να απαιτήσουν από τη δοσοληψία j χρηματικό ποσό μεγαλύτερο της τιμής που πρόσφεραν για την εξυπηρέτησή της πριν ληφθεί η απόφαση δρομολόγησης (price guarantees). Η οικονομική πολιτική που ακολουθείται και στους τρεις αλγόριθμους (διαμόρφωση των τιμών των επεξεργαστών) οδηγεί τελικά το σύστημα σε κατάσταση ισορροπίας (equilibrium state) κατά την οποία σε κάθε επεξεργαστή η ζήτηση ισούται με την προσφορά.

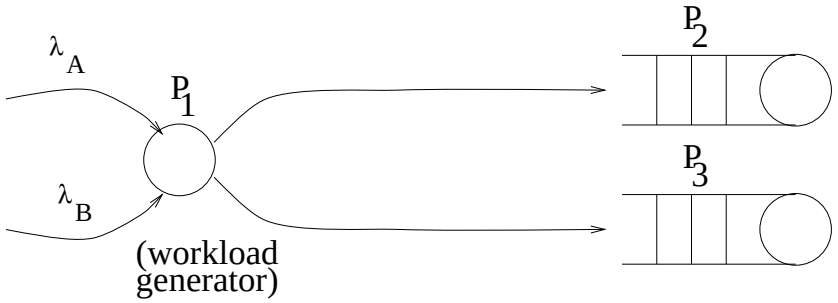
3.3 Μοντέλα Προσομοίωσης

3.3.1 Μοντέλα με μία πηγή φόρτου

Οι οικονομικοί αλγόριθμοι ECO, COMM, SQRD αποτέλεσαν τους αλγορίθμους δρομολόγησης σε πειράματα προσομοίωσης συστημάτων επεξεργασίας δοσοληψιών που έγιναν χρησιμοποιώντας τον προσομοιωτή TPSim. Τα μοντέλα συστημάτων που σχεδιάστηκαν για τη διεξαγωγή των πειραμάτων αυτών περιγράφονται στη συνέχεια.

- **Μοντέλο Τριών Κόμβων**

  Το σύστημα επεξεργασίας δοσοληψιών που εικονίζεται στο Σχήμα 3.3 αποτελείται



Σχήμα 3.3: Μοντέλο Τριών Κόμβων

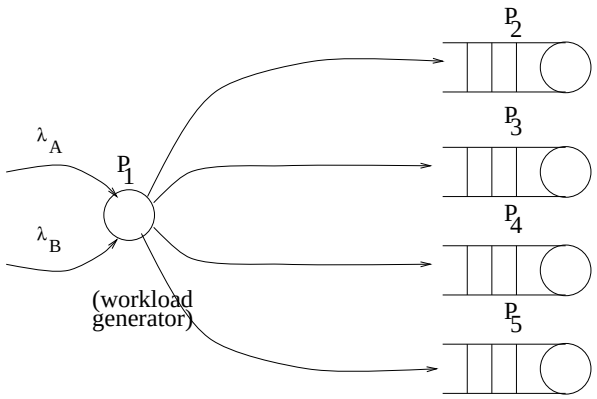
από τρεις κόμβους P_1, P_2, P_3. Ο κόμβος P_1 λειτουργεί ως ‘‘πηγή φόρτου’’ (frond-end) ενώ οι κόμβοι P_2, P_3 εκτελούν τις δοσοληψίες που υποβάλλονται στον κόμβο P_1

για εξυπηρέτηση από τερματικά (χρήστες). Χρησιμοποιείται το ανοικτό μοντέλο φόρτου εξυπηρέτησης. Οι κόμβοι P_2, P_3 έχουν από έναν επεξεργαστή των 55MIPS ο οποίος εφαρμόζει FirstComeFirstServed (FCFS) scheduling policy. Ο μέγιστος βαθμός πολυπρογραμματισμού (MPL) σε κάθε επεξεργαστή είναι 200. Ο ρυθμός μετάδοσης πακέτων (μεγέθους 1024 bytes) μεταξύ της πηγής φόρτου P_1 και των κόμβων εξυπηρέτησης P_2, P_3 είναι 4MBit/sec. Η βάση δεδομένων αποτελείται από σελίδες. Στους κόμβους P_2, P_3 τα δεδομένα κατανέμονται με τη round-robin μέθοδο σε δυο δίσκους δεδομένων. Το ημερολόγιο κάθε κόμβου (log) διατηρείται σε ξεχωριστό δίσκο. Οι δίσκοι είναι πανομοιότυποι και δίνουν μέσο χρόνο προσπέλασης περίπου ίσο με 10msec. Κάθε κόμβος εξυπηρέτησης διαθέτει ενταμιευτή δεδομένων μεγέθους 400.000 σελίδων. Το κόστος επικοινωνίας του κόμβου P_1 με τους κόμβους P_2, P_3 είναι 2.0 και 2.5 δρχ αντίστοιχα. Έχουν οριστεί δύο κλάσεις δοσοληψιών C_A, C_B. Σε όλες τις προσομοιώσεις στις οποίες θα χρησιμοποιηθεί το σύστημα αυτό οι δοσοληψίες των κλάσεων C_A, C_B έχουν ίση συχνότητα άφιξης στον κόμβο P_1 ($\lambda_A = \lambda_B$). Η κλάση C_A ζητά να προσπελάσει τη σχέση Α ενώ η κλάση C_B ζητά να προσπελάσει τη σχέση Β. Οι δοσοληψίες των κλάσεων C_A, C_B εκτελούν κατά μέσο όρο 10.000.000 εντολές (εκτός από τις εντολές για την εκτέλεση λειτουργιών του συστήματος) και από 1 έως 2 προσπελάσεις. Οι προσπελάσεις αυτές είναι πράξεις ανάγνωσης (read operations). Το μοντέλο αυτό προσομοιώθηκε για τρία διαφορετικά σχήματα ανάθεσης των σχέσεων της βάσης δεδομένων.

1. Πλήρη Αντιγραφή (Full Replication)
   Οι σχέσεις Α και Β έχουν αντιγραφεί σε όλους τους κόμβους.
2. Μερική Αντιγραφή (Half Replication)
   Η σχέση Α έχει αντιγραφεί στους κόμβους P_2, P_3 ενώ η σχέση Β έχει ανατεθεί μόνο στον κόμβο P_3.
3. Έλλειψη Αντιγραφής (No Replication)
   Οι σχέσεις Α και Β έχουν ανατεθεί στους κόμβους P_2 και P_3 αντίστοιχα.

- **Μοντέλο Πέντε Κόμβων**

Το σύστημα επεξεργασίας δοσοληψιών που εικονίζεται στο Σχήμα 3.4 αποτελείται από 5 κόμβους P_1, P_2, P_3, P_4, P_5. Ο κόμβος P_1 λειτουργεί ως ''πηγή φόρτου'' (frond-end) ενώ οι κόμβοι P_2, P_3, P_4, P_5 εκτελούν τις δοσοληψίες που υποβάλλονται στον κόμβο P_1



Σχήμα 3.4: Μοντέλο Πέντε Κόμβων

από ένα σύνολο τερματικών. Οι κόμβοι P_2, P_3, P_4, P_5 έχουν από ένα επεξεργαστή των 55MIPS ο οποίος εφαρμόζει FCFS scheduling policy. Η βάση δεδομένων αποτελείται από σελίδες. Κάθε κόμβος εξυπηρέτησης έχει ένα ενταμιευτή των 400.000 σελίδων. Το κόστος επικοινωνίας του κόμβου P_1 με τους κόμβους P_2, P_3, P_4, P_5 είναι 2.0, 2.5, 3.0, 2.0 δρχ αντίστοιχα Έχουν οριστεί δύο κλάσεις δοσοληψιών C_A, C_B. Η κλάση C_A ζητά να προσπελάσει τις σχέσεις Α, Β ενώ η κλάση C_B ζητά να προσπελάσει τις σχέσεις Γ, Δ. Οι δοσοληψίες των κλάσεων C_A, C_B εκτελούν κατά μέσο όρο 10.000.000 εντολές και από 1 έως 2 προσπελάσεις. Οι προσπελάσεις αυτές είναι πράξεις ανάγνωσης και αφορούν με την ίδια σχετική συχνότητα τις σχέσεις κάθε κλάσης. Οι δοσοληψίες των κλάσεων C_A, C_B έχουν τον ίδιο ρυθμό άφιξης στην ‘‘πηγή φόρτου’’ P_1. Το μοντέλο αυτό προσομοιώθηκε για τρία διαφορετικά σχήματα ανάθεσης των σχέσεων της βάσης δεδομένων.

1. Πλήρη Αντιγραφή (Full Replication)

   Οι σχέσεις Α, Β, Γ, Δ έχουν αντιγραφεί σε όλους τους κόμβους του συστήματος.

2. Μερική Αντιγραφή (Half Replication)

   Οι σχέσεις Α, Γ έχουν αντιγραφεί σε όλους τους κόμβους του συστήματος ενώ η σχέση Β έχει ανατεθεί μόνο στον κόμβο P_3 και η σχέση Δ έχει ανατεθεί μόνο στον κόμβο P_5.

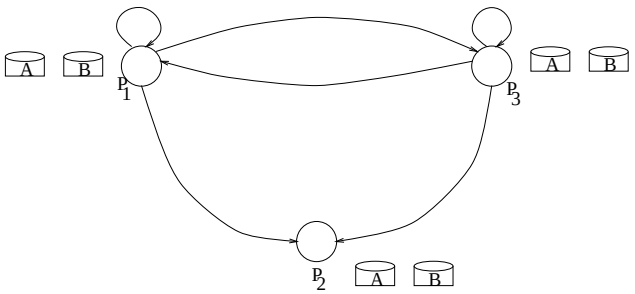
3. Έλλειψη Αντιγραφής (No Replication)

   Οι σχέσεις Α, Β, Γ, Δ έχουν ανατεθεί στους κόμβους P_2, P_3, P_4, P_5 αντίστοιχα.

3.3.2 Μοντέλα με πολλαπλές πηγές φόρτου

- **Μοντέλο με δύο πηγές φόρτου**

  Το σύστημα επεξεργασίας δοσοληψιών που εικονίζεται στο Σχήμα 3.5 αποτελείται

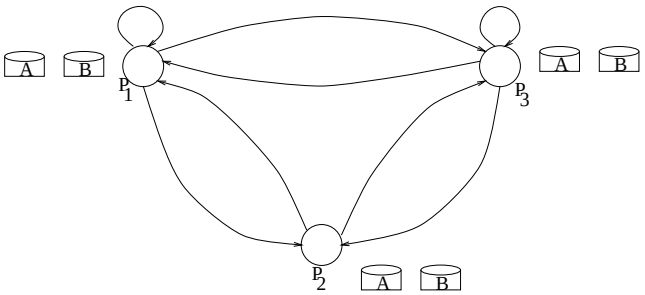


Σχήμα 3.5: Μοντέλο με δύο πηγές φόρτου

από 3 κόμβους P_1, P_2, P_3. Στους κόμβους P_1, P_3 υποβάλλονται, από ένα σύνολο τερματικών, δοσοληψίες οι οποίες είναι δυνατό να εξυπηρετηθούν σε όλους τους κόμβους του συστήματος. Κάθε κόμβος έχει ένα επεξεργαστή των 55MIPS ο οποίος εφαρμόζει FCFS scheduling policy. Η βάση δεδομένων αποτελείται από σελίδες. Ο ενταμιευτής κάθε κόμβου έχει ένα μέγεθος 400.000 σελίδες. Το κόστος επικοινωνίας του κόμβου P_1 με τους κόμβους P_2, P_3 είναι 2.0 και 2.5 δρχ αντίστοιχα. Το κόστος επικοινωνίας του κόμβου P_2 με τους κόμβους P_1, P_3 είναι 2.5 και 3.0 δρχ αντίστοιχα ενώ το κόστος επικοινωνίας του κόμβου P_3 με τους κόμβους P_1, P_2 είναι 2.0 και 2.5 δρχ αντίστοιχα. Έχουν οριστεί δύο κλάσεις δοσοληψιών C_A, C_B. Οι δοσοληψίες των κλάσεων Α, Β έχουν την ίδια συχνότητα άφιξης στους κόμβους P_1, P_3. Η κλάση C_A ζητά να προσπελάσει τη σχέση Α, ενώ η κλάση C_B ζητά να προσπελάσει τη σχέση Β. Οι δοσοληψίες των κλάσεων C_A, C_B εκτελούν κατά μέσο όρο 10.000.000 εντολές και μία προσπέλαση/ανάγνωση στις σχέσεις Α, Β αντίστοιχα. Οι σχέσεις Α, Β έχουν αντιγραφεί σε όλους τους κόμβους του συστήματος (πλήρης αντιγραφή).

- **Μοντέλο με τρείς πηγές φόρτου**

  Στο σύστημα επεξεργασίας δοσοληψιών που εικονίζεται στο Σχήμα 3.6 ο κόμβος P_2 αποτελεί και αυτός ‘‘πηγή φόρτου’’. Αυτή είναι η μόνη διαφορά του μοντέλου αυτού από το προηγούμενο. Στον κόμβο P_2 υποβάλλονται δοσοληψίες από ένα σύνολο τερματικών. Ο ρυθμός άφιξης δοσοληψιών των κλάσεων C_A, C_B είναι ίδιος σε όλους

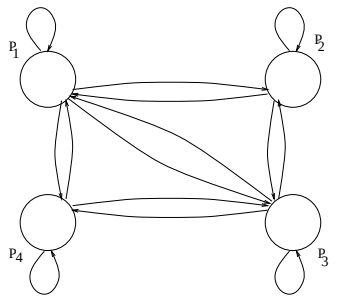


Σχήμα 3.6: Μοντέλο με τρείς πηγές φόρτου

τους κόμβους του συστήματος.

- **Μοντέλο με τέσσερεις πηγές φόρτου**

  Το σύστημα επεξεργασίας δοσοληψιών που εικονίζεται στο Σχήμα 3.7 αποτελείται



Σχήμα 3.7: Μοντέλο με τέσσερεις πηγές φόρτου

από 4 κόμβους P_1, P_2, P_3, P_4. Σε κάθε κόμβο υποβάλλονται δοσοληψίες από 40 τερματικά. Κάθε κόμβος έχει ένα επεξεργαστή των 55MIPS ο οποίος εφαρμόζει FCFS scheduling policy. Το κόστος επικοινωνίας του κόμβου P_1 με τους κόμβους P_2, P_3, P_4 είναι 2.0, 2.5 και 3.0 δρχ. αντίστοιχα. Το κόστος επικοινωνίας του κόμβου P_2 με τους κόμβους P_1, P_3, P_4 είναι 2.5, 3.0 και 2.5 δρχ. αντίστοιχα. Το κόστος επικοινωνίας του κόμβου P_3 με τους κόμβους P_1, P_2, P_4 είναι 2.0, 2.5 και 2.0 δρχ. ενώ το κόστος επικοινωνίας του κόμβου P_4 με τους κόμβους P_1, P_2, P_3 είναι είναι 2.0, 2.5 και 3.0 δρχ. αντίστοιχα. Έχουν οριστεί 4 κλάσεις δοσοληψιών $C_A, C_B, C_\Gamma, C_\Delta$ που επιθυμούν να προσπελάσουν τις σχέσεις Α, Β, Γ, Δ αντίστοιχα. Οι δοσοληψίες των κλάσεων $C_A, C_B, C_\Gamma, C_\Delta$ εκτελούν κατά μέσο όρο 10.000.000 εντολές και μία

προσπέλαση/ανάγνωση στις σχέσεις Α, Β Γ, Δ αντίστοιχα. Οι σχέσεις Α, Β Γ, Δ έχουν ανατεθεί σε όλους τους κόμβους P_1, P_2, P_3, P_4.

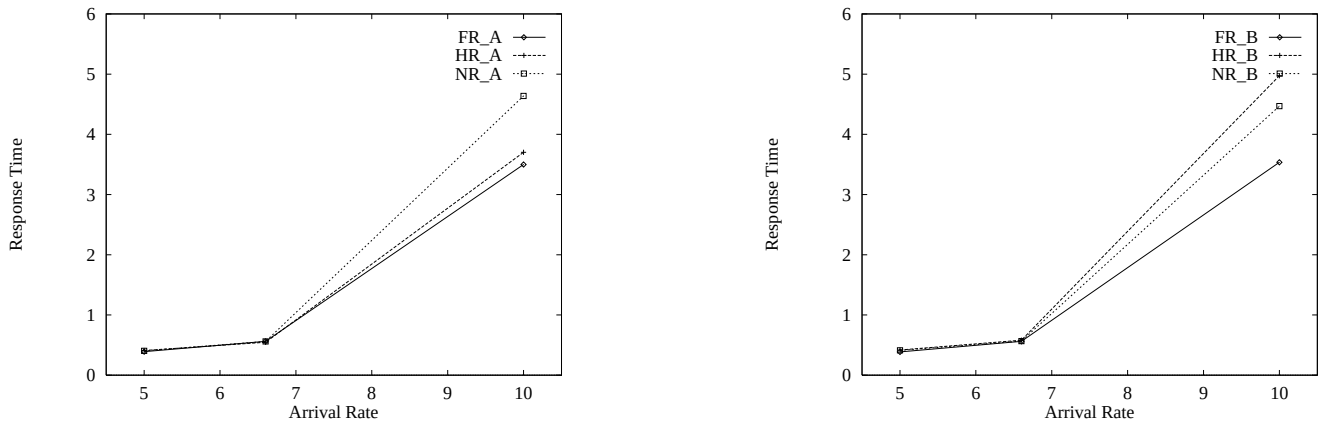
3.4 Αποτελέσματα

Οι πρώτες προσομοιώσεις πραγματοποιήθηκαν για να εντοπιστεί για κάθε οικονομικό αλγόριθμο σε διαφορετικές συνθήκες φόρτου του συστήματος ποιό από τα προτεινόμενα σχήματα ανάθεσης επιτυγχάνει ελαχιστοποίηση του μέσου χρόνου απόκρισης κάθε κλάσης. Στις γραφικές παραστάσεις που ακολουθούν στον άξονα Χ καταγράφεται ο ρυθμός άφιξης της κλάσης ενώ στον άξονα Υ ο μέσος χρόνος απόκρισης της κλάσης. Κάθε σημείο μιας γραφικής παράστασης καταγράφει για συγκεκριμένο φόρτο του συστήματος επεξεργασίας δοσοληψιών το μέσο χρόνο απόκρισης μιας κλάσης του συστήματος. Στο σύστημα επεξεργασίας δοσοληψιών ακολουθείται ένα από τα τρία σχήματα ανάθεσης δεδομένων που έχουν οριστεί και η δρομολόγηση γίνεται με ένα από τρεις οικονομικούς αλγορίθμους κατανομής φόρτου.

Στο Σχήμα 3.8 απεικονίζεται η συμπεριφορά του μοντέλου επεξεργασίας δοσοληψιών τριών κόμβων όταν η δρομολόγηση γίνεται με βάση τον οικονομικό αλγόριθμο ECO, στο Σχήμα 3.9 με βάση τον αλγόριθμο COMM ενώ στο Σχήμα 3.10 με βάση τον οικονομικό αλγόριθμο SQRD. Κάθε σχήμα περιλαμβάνει 3 γραφικές παραστάσεις που αντιστοιχούν στα 3 διαφορετικά σχήματα ανάθεσης.

Στο Σχήμα 3.11 απεικονίζεται η συμπεριφορά του μοντέλου επεξεργασίας δοσοληψιών πέντε κόμβων όταν ο αλγόριθμος κατανομής φόρτου είναι ο ECO, στο Σχήμα 3.12 όταν ο αλγόριθμος είναι ο COMM ενώ στο Σχήμα 3.13 όταν η δρομολόγηση γίνεται με τον οικονομικό αλγόριθμο SQRD.

Οι γραφικές παραστάσεις δείχνουν ότι στην πλειοψηφία των πειραμάτων η αντιγραφή των σχέσεων της βάσης σε όλους τους επεξεργαστές του συστήματος δίνει ελάχιστο χρόνο απόκρισης για τις κλάσεις C_A, C_B ενώ η ανάθεση κάθε σχέσης σε ένα μόνο επεξεργαστή δίνει μέγιστο μέσο χρόνο απόκρισης. Στην πλειοψηφία των πειραμάτων ο μέσος χρόνος απόκρισης των δοσοληψιών κάθε κλάσης που επιτυγχάνεται με μερική αντιγραφή των σχέσεων της βάσης ανήκει στο διάστημα που ορίζουν οι μέσοι χρόνοι απόκρισης κάθε κλάσης που επιτυγχάνονται με πλήρη αντιγραφή και μοναδική ανάθεση των σχέσεων της βάσης. Στην περίπτωση της μερικής αντιγραφής οι δύο κλάσεις του

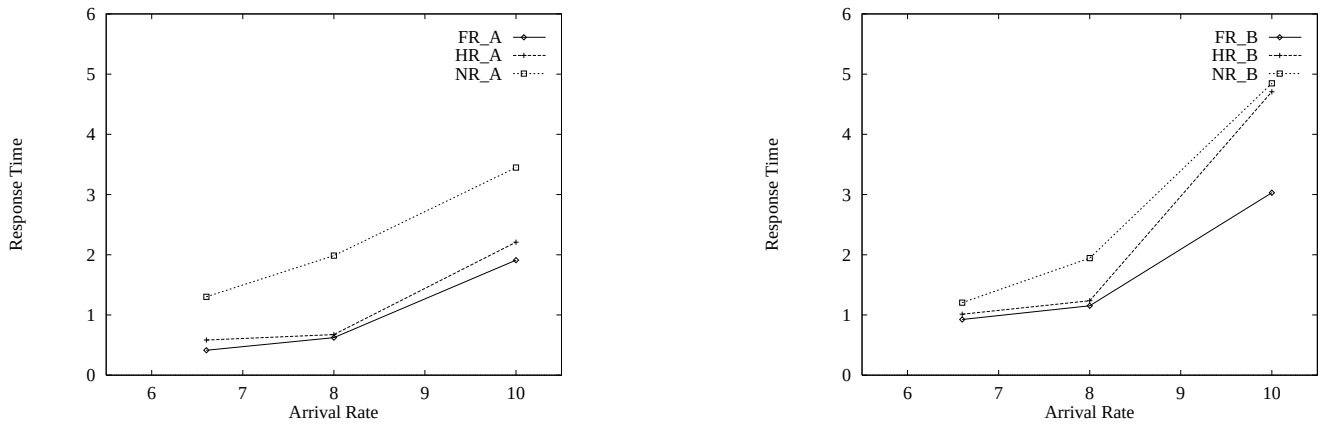


Σχήμα 3.8:

Αποτελέσματα προσομοίωσης του Μοντέλου Τριών Κόμβων χρησιμοποιώντας τον αλγόριθμο δρομολόγησης ECO. Η μέση χρησιμοποίηση των κόμβων του συστήματος για ρυθμό άφιξης 5 transactions/sec είναι 48%, ενώ για ρυθμό άφιξης 6.6 είναι 64% και για ρυθμό άφιξης 10 είναι 94%.

συστήματος ανταγωνίζονται για το αγαθό (υπολογιστικό χρόνο) στους κόμβους στους οποίους υπάρχουν όλα τα δεδομένα και συνήθως μια μόνο κλάση επωφελείται. Έτσι υπάρχουν πειράματα στα οποία για μία από τις δύο κλάσεις C_j ο μέσος χρόνος απόκρισης για μερική αντιγραφή των σχέσεων είναι μεγαλύτερος από το μέσο χρόνο απόκρισης της κλάσης C_j για μη αντιγραφή των σχέσεων ή μικρότερος από το μέσο χρόνο απόκρισης της κλάσης C_j για πλήρη αντιγραφή των σχέσεων. Οι γράφοι 3.8 και 3.13 δείχνουν ανάλογη συμπεριφορά.

Στο σύνολο των προηγούμενων προσομοιώσεων οι κλάσεις C_A, C_B είχαν τον ίδιο ρυθμό άφιξης (λ_A, λ_B) και το ίδιο κεφάλαιο ($budget_A = budget_B$). Διεξήχθησαν πειράματα στα οποία οι κλάσεις είχαν τον ίδιο ρυθμό άφιξης αλλά διαφορετικά κεφάλαια ($budget_A \neq budget_B$). Ουσιαστικά το κεφάλαιο αποτελεί μέτρο της προτεραιότητας κάθε κλάσης αφού η κλάση με σχετικά μεγαλύτερο κεφάλαιο μπορεί να διεκδικήσει περισσότερα αγαθά ή ποσοστό ενός αγαθού και να επιτύχει καλύτερη απόδοση (μικρότερο μέσο χρόνο απόκρισης). Επίσης όσο μεγαλύτερο είναι το κεφάλαιο μιας κλάσης (προτεραιότητα) τόσο μικρότερη είναι η πιθανότητα το κεφάλαιο αυτό να μην επαρκεί για την αγορά των απαραίτητων αγαθών (πόρων). Αν το κεφάλαιο δεν επαρκεί (μικρή προτεραιότητα) η δοσοληψία απορρίπτεται από το σύστημα.



Σχήμα 3.9:

Αποτελέσματα προσομοίωσης του Μοντέλου Τριών Κόμβων χρησιμοποιώντας τον αλγόριθμο δρομολόγησης COMM. Η μέση χρησιμοποίηση των κόμβων του συστήματος για ρυθμό άφιξης 6.6 transactions/sec είναι 63%, ενώ για ρυθμό άφιξης 8 είναι 76% και για ρυθμό άφιξης 10 είναι 94%

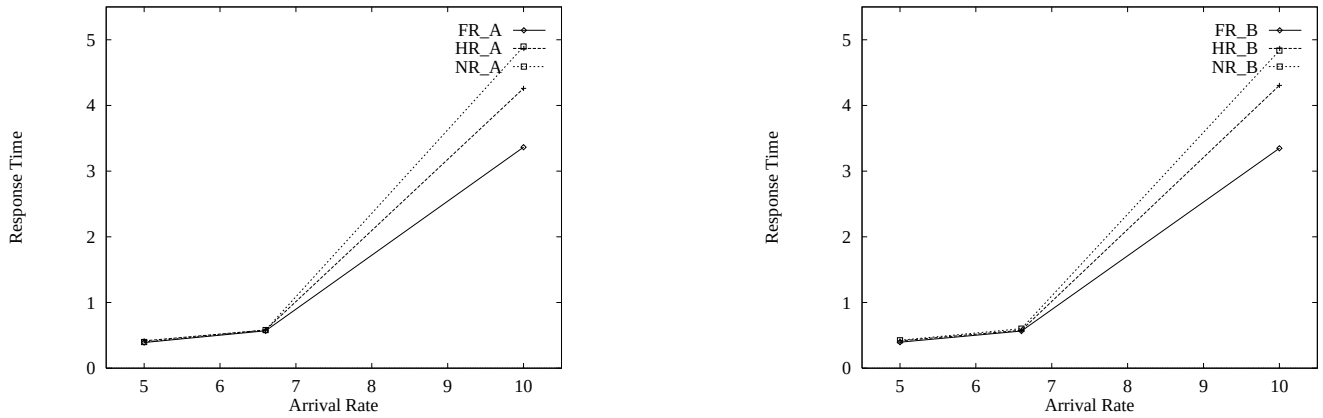
Budget A	Budget B	Utility 1	Utility 2	Resp A	Resp B	Aborts A	Aborts B
850	850	0.96	0.97	2.41	2.43	0%	0%
8.85	850	0.85	0.87	2.39	2.00	20.5%	0%
7.0	850	0.72	0.72	2.10	1.25	50%	0%
8.75	8.75	0.61	0.60	2.14	2.21	36%	36%

Πίνακας 1

Στον Πίνακα 1 παρουσιάζονται τα αποτελέσματα κάποιων προσομοιώσεων που έγιναν χρησιμοποιώντας το μοντέλο επεξεργασίας δοσοληψιών τριών κόμβων του Σχήματος 3.3 για πλήρη αντιγραφή των σχέσεων της βάσης. Ο οικονομικός αλγόριθμος ECO χρησιμοποιείται για δρομολόγηση των δοσοληψιών του συστήματος.

Παρατηρούμε ότι όταν οι κλάσεις C_A, C_B διαθέτουν κεφάλαιο ίσο με 850 δρχ. αντίστοιχα επιτυγχάνουν εξυπηρέτηση των δοσοληψιών τους με μέσο χρόνο απόκρισης 2.41 και 2.43 seconds αντίστοιχα. Η πιθανότητα να απορριφθεί κάποια δοσοληψία από το σύστημα λόγω έλλειψης του αναγκαίου για εξυπηρέτηση κεφαλαίου είναι ίση με 0 και για τις δύο κλάσεις.

Αν το κεφάλαιο της κλάσης C_A μειωθεί σε 8.85 δρχ. ένα ποσοστό 20% των δοσοληψιών της κλάσης αυτής θα απορριφθεί από το σύστημα. Η κλάση C_B η οποία διαθέτει



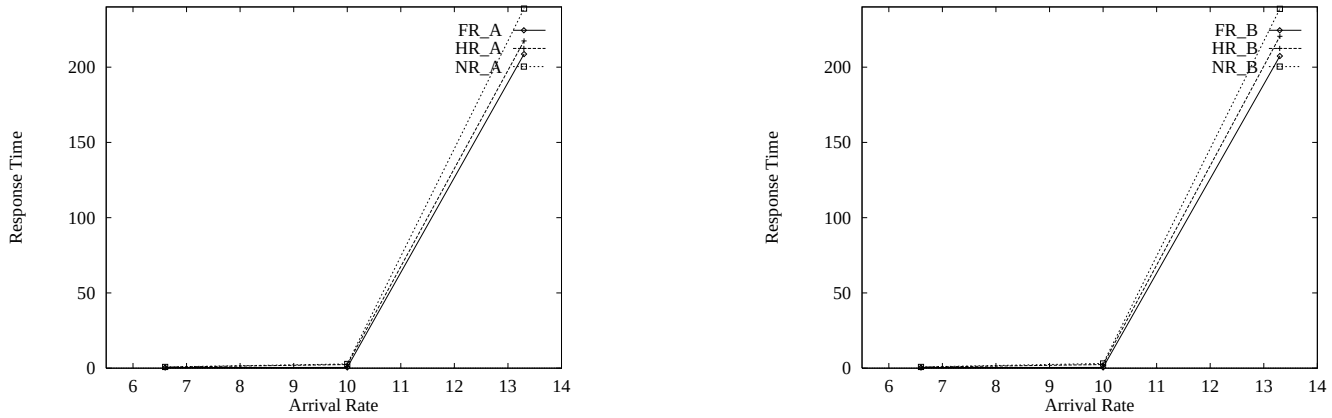
Σχήμα 3.10:

Αποτελέσματα προσομοίωσης του Μοντέλου Τριών Κόμβων χρησιμοποιώντας τον αλγόριθμο δρομολόγησης SQRD. Η μέση χρησιμοποίηση των κόμβων του συστήματος για ρυθμό άφιξης 5 transactions/sec είναι 48%, ενώ για ρυθμό άφιξης 6.6 είναι 64% και για ρυθμό άφιξης 10 είναι 94%.

μεγαλύτερο κεφάλαιο θα επωφεληθεί, θα αγοράσει μεγαλύτερο ποσοστό του αγαθού (υπολογιστικό χρόνο) και θα μειώσει το μέσο χρόνο που απαιτείται για εξυπηρέτηση των δοσοληψιών της από 2.43 σε 2.00 seconds. Καμιά δοσοληψία της κλάσης C_B δε θα απορριφθεί. Η κλάση C_A θα μειώσει και αυτή το μέσο χρόνο απόκρισης των δοσοληψιών της αφού στο σύστημα υπάρχουν πλέον λιγότερες δοσοληψίες της κλάσης αυτής που επιθυμούν (και έχουν την δυνατότητα) να εξυπηρετηθούν.

Αν το κεφάλαιο της κλάσης C_A μειωθεί ακόμη περισσότερο (7 δρχ.) ένα ποσοστό 50% των δοσοληψιών της κλάσης θα απορριφθεί από το σύστημα. Η κλάση C_B θα μειώσει σημαντικά το μέσο χρόνο που απαιτείται για εξυπηρέτηση των δοσοληψιών της (1.25 seconds) αφού διαθέτει ακόμη μεγαλύτερο κεφάλαιο σε σχέση με την κλάση C_A και έτσι μπορεί να αγοράσει ακόμη περισσότερο υπολογιστικό χρόνο.

Αν μειωθεί σε 8.75 δρχ. το κεφάλαιο τόσο της κλάσης C_A όσο και της κλάσης C_B ένα ποσοστό 36% των δοσοληψιών κάθε κλάσης θα απορριφθεί από το σύστημα. Οι μέσοι χρόνοι απόκρισης των κλάσεων C_A, C_B θα μειωθούν από 2.41, 2.43 seconds που ήταν όταν οι κλάσεις διέθεταν επαρκή κεφάλαια σε 2.14, 2.24 seconds αντίστοιχα. Η μείωση οφείλεται στην ελάττωση του αριθμού των δοσοληψιών κάθε κλάσης που εξυπηρετούνται από το σύστημα.



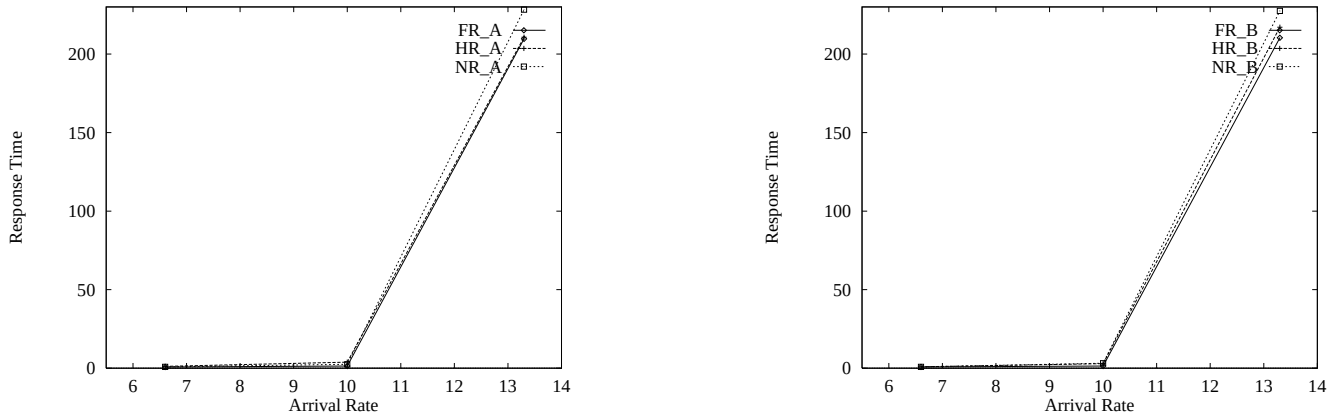
Σχήμα 3.11:

Αποτελέσματα προσομοίωσης του Μοντέλου Πέντε Κόμβων χρησιμοποιώντας τον αλγόριθμο δρομολόγησης ECO. Η μέση χρησιμοποίηση των κόμβων του συστήματος για ρυθμό άφιξης 6.6 transactions/sec είναι 57%, ενώ για ρυθμό άφιξης 10 είναι 86% και για ρυθμό άφιξης 13.3 είναι 99%.

Η κλάση C_B παρουσιάζει στο πείραμα αυτό μεγαλύτερο μέσο χρόνο απόκρισης των δοσοληψιών της (2.21 seconds) σε σχέση με τα προηγούμενα πειράματα στα οποία όμως το κεφάλαιό της δεν είχε υποστεί μείωση. Αυτό οφείλεται στο ότι δε διαθέτει πλέον ισχυρότερο κεφάλαιο από την κλάση C_A ώστε να διεκδικήσει μεγαλύτερο ποσοστό του υπολογιστικού χρόνου. Η χρησιμοποίηση των κόμβων σε κάθε πείραμα είναι ανάλογη του πλήθους των δοσοληψιών που εξυπηρετούνται από το σύστημα. Όσο μεγαλύτερο είναι το ποσοστό των δοσοληψιών που απορρίπτονται από το σύστημα λόγω ανεπάρκειας του κεφαλαίου τόσο μικρότερη είναι η χρησιμοποίηση (utilization) των κόμβων του συστήματος. Το ότι οι κόμβοι P_2, P_3 παρουσιάζουν σε όλα τα πειράματα το ίδιο ποσοστό χρησιμοποίησης αποδεικνύει ότι ο οικονομικός αλγόριθμος ECO ισοκατανέμει το φόρτο των κλάσεων C_A, C_B στους επεξεργαστές.

Στο σύνολο των προηγούμενων προσομοιώσεων υπάρχει μια μόνο πηγή φόρτου. Έγιναν προσομοιώσεις για να μελετηθεί η συμπεριφορά των αλγορίθμων σε συστήματα επεξεργασίας δοσοληψιών με περισσότερες πηγές φόρτου.

	Utility 1	Utility 2	Utility 3	Resp A	Resp B
2 πηγές φόρτου	0.76	0.74	0.51	0.32	0.31
3 πηγές φόρτου	0.97	0.95	0.93	1.64	1.65



Σχήμα 3.12:

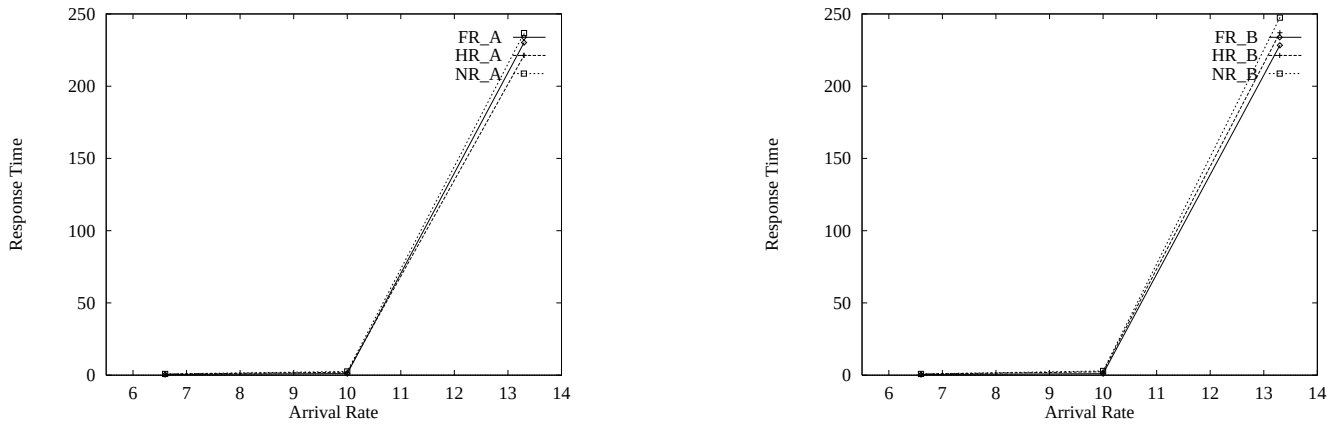
Αποτελέσματα προσομοίωσης του Μοντέλου Πέντε Κόμβων χρησιμοποιώντας τον αλγόριθμο δρομολόγησης COMM. Η μέση χρησιμοποίηση των κόμβων του συστήματος για ρυθμό άφιξης 6.6 transactions/sec είναι 56%, ενώ για ρυθμό άφιξης 10 είναι 86% και για ρυθμό άφιξης 13.3 είναι 99%.

Πίνακας 2

Στην πρώτη γραμμή του Πίνακα 2 παραθέτονται τα αποτελέσματα προσομοίωσης του μοντέλου επεξεργασίας δοσοληψιών με δύο πηγές φόρτου του Σχήματος 3.5 ενώ στη δεύτερη γραμμή τα αποτελέσματα προσομοίωσης του μοντέλου με τρείς πηγές φόρτου του Σχήματος 3.6. Και στις δύο προσομοιώσεις η δρομολόγηση δοσοληψιών έγινε με βάση τον αλγόριθμο SQRD.

Ο αλγόριθμος SQRD επιλέγει να δρομολογήσει μια δοσοληψία στον κόμβο που προσφέρει μικρότερη τιμή για τις υπηρεσίες του. (τιμή = (κόστος απόκρισης)^2 + κόστος μεταβίβασης). Το κόστος μεταβίβασης της δοσοληψίας από την πηγή φόρτου στον πιθανό επεξεργαστή είναι μηδενικό όταν η δοσοληψία εκτελείται τοπικά. Οι κόμβοι P_1, P_3 (πηγές φόρτου) παρουσιάζουν επομένως πλεονέκτημα αφού δε λαμβάνουν υπόψη το κόστος μεταβίβασης στη διαμόρφωση της τιμής των υπηρεσιών τους. Έχουν λοιπόν μεγαλύτερη πιθανότητα να εξυπηρετήσουν μια δοσοληψία σε σχέση με τον κόμβο P_2. Γι'αυτό το λόγο όταν υπάρχουν δύο πηγές φόρτου P_1, P_3 το ποσοστό χρησιμοποίησης τους είναι μεγαλύτερο του ποσοστού χρησιμοποίησης του κόμβου P_2 (παρά το γεγονός ότι όλοι οι κόμβοι έχουν την ίδια υπολογιστική δύναμη και διαθέτουν τα ίδια δεδομένα).

Αντίθετα όταν υπάρχουν 3 πηγές φόρτου οι επεξεργαστές του συστήματος



Σχήμα 3.13:

Αποτελέσματα προσομοίωσης του Μοντέλου Πέντε Κόμβων χρησιμοποιώντας τον αλγόριθμο δρομολόγησης SQRD. Η μέση χρησιμοποίηση των κόμβων του συστήματος για ρυθμό άφιξης 6.6 transactions/sec είναι 58%, ενώ για ρυθμό άφιξης 10 είναι 85% και για ρυθμό άφιξης 13.3 είναι 99%.

παρουσιάζουν το ίδιο ποσοστό χρησιμοποίησης. Όταν υπάρχουν 3 πηγές φόρτου ο μέσος χρόνος απόκρισης των δοσοληψιών των δύο κλάσεων αυξάνει αφού αυξάνει ο φόρτος του συστήματος.

Στα πειράματα που παρατέθηκαν μέχρι τώρα ο ρυθμός άφιξης των δοσοληψιών κάθε κλάσης διατηρείται σταθερός σε όλη τη διάρκεια της προσομοίωσης. Το μοντέλο επεξεργασίας δοσοληψιών με τέσσερεις πηγές φόρτου του Σχήματος 3.7 χρησιμοποιήθηκε σε προσομοιώσεις στις οποίες ο χρόνος προσομοίωσης διαιρείται σε διαστήματα στα οποία παρουσιάζεται μεταβολή (αύξηση) του ρυθμού άφιξης των δοσοληψιών. Σε όλες τις προσομοιώσεις υπάρχει πλήρης αντιγραφή των σχέσεων της βάσης στους κόμβους του συστήματος. Οι κλάσεις δοσοληψιών $C_A, C_\mathrm{B}, C_\Gamma, C_\Delta$ έχουν σε κάθε διάστημα του χρόνου προσομοίωσης ίσο μεταξύ τους φόρτο.

	Resp A	Resp B	Resp Γ	Resp Δ
ECO	0.54	0.55	0.55	0.55
COMM	1.24	1.24	1.26	1.25
SQRD	0.52	0.52	0.52	0.55

Πίνακας 3

Παρατηρούμε στον Πίνακα 3 ότι οι κλάσεις μεταξύ τους παρουσιάζουν για κάθε

οικονομικό αλγόριθμο δρομολόγησης παρόμοιο μέσο χρόνο απόκρισης.

Οι αλγόριθμοι ECO, SQRD, στους οποίους οι τιμές των κόμβων διαμορφώνονται δίνοντας ιδιαίτερο βάρος στον αναμενόμενο χρόνο απόκρισης των δοσοληψιών παρουσιάζουν καλύτερη απόδοση (μικρότερο μέσο χρόνο απόκρισης για κάθε κλάση δοσοληψιών).

Αντίθετα ο αλγόριθμος COMM παρουσιάζει χειρότερα αποτελέσματα. Αυτό συμβαίνει γιατί όταν εφαρμόζεται ο αλγόριθμος COMM οι τιμές των κόμβων του συστήματος διαμορφώνονται δίνοντας την ίδια βαρύτητα στον αναμενόμενο χρόνο απόκρισης και στο κόστος μεταβίβασης μιας δοσοληψίας. Στο σύστημα του Σχήματος 3.7 ένα μεγάλο ποσοστό των δοσοληψιών που παράγονται σε ένα κόμβο επιλέγουν να εξυπηρετηθούν τοπικά όταν εφαρμόζεται ο αλγόριθμος δρομολόγησης COMM. Έτσι οι δοσοληψίες επιτυγχάνουν να μειώσουν τα έξοδα τους αφού δε χρεώνονται τη μεταβίβασή τους σε άλλο κόμβο παρά το γεγονός ότι ο κόμβος αυτός δύναται να παρουσιάζει φόρτο μικρότερο από το δικό τους (και έτσι μπορεί να εγγυηθεί μικρότερο μέσο χρόνο απόκρισης).

Συμπεραίνουμε λοιπόν ότι στην κοστολόγηση των υπηρεσιών των κόμβων αυτών των συστημάτων είναι αποδοτικότερο να δίνεται μεγαλύτερο βάρος στον αναμενόμενο χρόνο απόκρισης των δοσοληψιών.

	Resp C_A	Resp C_B	Resp C_Γ	Resp C_Δ
ECO	0.31	0.31	0.31	0.31
COMM	0.64	0.64	0.63	0.64
SQRD	0.31	0.31	0.30	0.30

Πίνακας 4

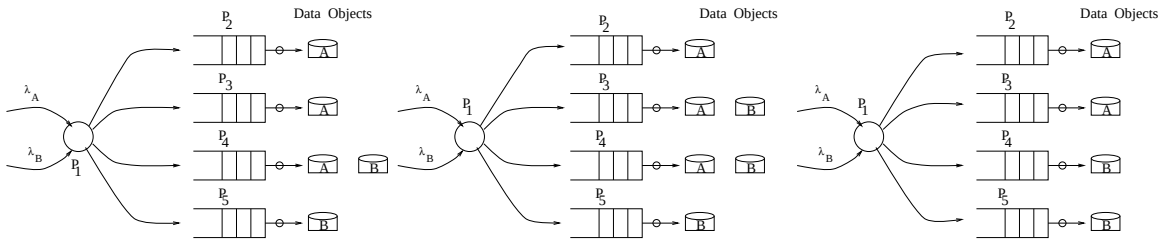
Στον Πίνακα 4 παραθέτονται τα αποτελέσματα προσομοιώσεων στις οποίες χρησιμοποιήθηκε μοντέλο συστήματος ίδιο με το προηγούμενο (μοντέλο με τέσσερεις πηγές φόρτου). Η μόνη διαφορά είναι ότι σε κάθε διάστημα του χρόνου προσομοίωσης οι κλάσεις C_A, C_B παρουσιάζουν μικρότερο φόρτο από τις κλάσεις C_Γ, C_Δ. Η διαφορά που παρουσιάζει ο Πίνακας 4 από τον Πίνακα 3 είναι η μείωση του μέσου χρόνου απόκρισης κάθε κλάσης. Αυτό οφείλεται στην ελάττωση του αριθμού των δοσοληψιών στο σύστημα. Η διαφορά φόρτου που παρουσιάζουν οι κλάσεις δεν επηρεάζει το μέσο χρόνο απόκρισης δοσοληψιών κάθε κλάσης.

3.5 Ανάλυση

	Πρώτο Μοντέλο	Δεύτερο Μοντέλο	Τρίτο Μοντέλο
Κλάση C_A	2.26	1.99	2.26
Κλάση C_B	2.76	2.29	2.26

Πίνακας 5

Στον Πίνακα 5 παραθέτονται τα αποτελέσματα προσομοίωσης τριών συστημάτων επεξεργασίας δοσοληψιών στα οποία η δρομολόγηση εργασιών γίνεται με βάση τον οικονομικό αλγόριθμο ECO. Τα συστήματα αυτά εικονίζονται στο Σχήμα 3.5.



Πρώτο Μοντέλο Δεύτερο Μοντέλο Τρίτο Μοντέλο

Σχήμα 3.14: Μοντέλα Ανάλυσης

Τα συστήματα διαφέρουν μόνο στην ανάθεση των σχέσεων της βάσης στους κόμβους. Κάθε σύστημα αποτελείται από 5 κόμβους P_1, P_2, P_3, P_4, P_5. Ο κόμβος P_1 λειτουργεί ως πηγή φόρτου ενώ οι κόμβοι P_2, P_3, P_4, P_5 εκτελούν δοσοληψίες. Κάθε κόμβος έχει ένα επεξεργαστή 55MIPS. Έχουν οριστεί δύο κλάσεις δοσοληψιών $C_A, C_{\rm B}$. Η κλάση C_A ζητά να προσπελάσει τη σχέση A ενώ η κλάση $C_{\rm B}$ ζητά να προσπελάσει τη σχέση B. Παρατηρούμε ότι μικρότερο μέσο χρόνο απόκρισης παρουσιάζουν οι δοσοληψίες που εξυπηρετούνται στο σύστημα 2 (ανεξάρτητα από την κλάση στην οποία ανήκουν).

Η ανάλυση των τριών συστημάτων επεξεργασίας δοσοληψιών με βάση τη θεωρία ουρών αποδεικνύει την ορθότητα των αποτελεσμάτων των προσομοιώσεων. Κάθε κόμβος των συστημάτων αυτών έχει έναν επεξεργαστή και μία ουρά αναμονής για τις δοσοληψίες που του ανατείθονται. Ο ρυθμός άφιξης δοσοληψιών στη πηγή φόρτου ακολουθεί Poisson κατανομή. Ο παραγόμενος φόρτος διασπάται σε επιμέρους κατανομές Poisson πριν ανατεθεί στους επεξεργαστές του συστήματος. Διαπιστώνουμε λοιπόν οτι οι ουρές

αναμονής των κόμβων συμπεριφέρονται ως M/M/1 ουρές. Με βάση τη θεωρία ουρών το πρόβλημα της κατανομής του φόρτου σε κάθε ένα από αυτά τα συστήματα ορίζεται ως εξής :

min T
given λ_A, λ_B
μ_A, μ_B

constraints $\lambda_A = \sum_{i=1}^{4} \lambda_{Ai}$
$\lambda_B = \sum_{i=1}^{4} \lambda_{Bi}$
$i \in [1, 4]$

Όπου T είναι ο μέσος χρόνος απόκρισης των δοσοληψιών λ_A, λ_B είναι οι ρυθμοί άφιξης των κλάσεων C_A, C_B αντίστοιχα ενώ μ_A, μ_B είναι οι χρόνοι εξυπηρέτησης των δοσοληψιών των κλάσεων C_A, C_B αντίστοιχα σε ιδανικό επεξεργαστή $P_i(r_i = 1)$. Με λ_{Ai} συμβολίζεται ο ρυθμός άφιξης δοσοληψιών της κλάσης C_A στον κόμβο i όπου $i \in [1, 4]$. Αναζητείται λοιπόν διάνυσμα ανάθεσης φόρτου στους 4 επεξεργαστές $(\lambda^*_{A1}, \lambda^*_{A2}, \lambda^*_{A3}, \lambda^*_{A4}, \lambda^*_{B1}, \lambda^*_{B2}, \lambda^*_{B3}, \lambda^*_{B4})$ το οποίο να δίνει ελάχιστο μέσο φόρτο απόκρισης των δοσοληψιών του συστήματος.

Για το Πρώτο Μοντέλο έχουμε

$$\lambda_A = \lambda_{A1} + \lambda_{A2} + \lambda_{A3}$$

$$\lambda_B = \lambda_{B3} + \lambda_{B4}$$

$$N_1 = \frac{\lambda_{A1}}{\mu_A - \lambda_{A1}} + \frac{\lambda_{A2}}{\mu_A - \lambda_{A2}} + \frac{\lambda_{A3} + \lambda_{B3}}{\frac{\lambda_{A3} \cdot \mu_A + \lambda_{B3} \cdot \mu_B}{\lambda_{A3} + \lambda_{B3}} - (\lambda_{A3} + \lambda_{B3})} + \frac{\lambda_{B4}}{\mu_B - \lambda_{B4}}$$

όπου N_1 είναι ο αναμενόμενος αριθμός δοσοληψιών στο σύστημα ενώ ο μέσος χρόνος απόκρισης των δοσοληψιών του συστήματος είναι

$$T_1 = \frac{N_1}{\lambda_A + \lambda_B}$$

Ανάλογα για το Δεύτερο Μοντέλο έχουμε

$$\lambda_A = \lambda_{A1} + \lambda_{A2} + \lambda_{A3}$$

$$\lambda_B = \lambda_{B2} + \lambda_{B3} + \lambda_{B4}$$

$$N_2 = \frac{\lambda_{A1}}{\mu_A - \lambda_{A1}} + \frac{\lambda_{A2} + \lambda_{B2}}{\frac{\lambda_{A2}\cdot\mu_A + \lambda_{B2}\cdot\mu_B}{\lambda_{A2}+\lambda_{B2}} - (\lambda_{A2} + \lambda_{B2})} + \frac{\lambda_{A3} + \lambda_{B3}}{\frac{\lambda_{A3}\cdot\mu_A + \lambda_{B3}\cdot\mu_B}{\lambda_{A3}+\lambda_{B3}} - (\lambda_{A3} + \lambda_{B3})} + \frac{\lambda_{B4}}{\mu_B - \lambda_{B4}}$$

$$T_2 = \frac{N_2}{\lambda_A + \lambda_B}$$

ενώ για το Τρίτο Μοντέλο έχουμε

$$\lambda_A = \lambda_{A1} + \lambda_{A2}$$

$$\lambda_B = \lambda_{B3} + \lambda_{B4}$$

$$N_3 = \frac{\lambda_{A1}}{\mu_A - \lambda_{A1}} + \frac{\lambda_{A2}}{\mu_A - \lambda_{A2}} + \frac{\lambda_{B3}}{\mu_B - \lambda_{B3}} + \frac{\lambda_{A4}}{\mu_B - \lambda_{B4}}$$

$$T_3 = \frac{N_3}{\lambda_A + \lambda_B}$$

Τα τρία προβλήματα ελαχιστοποίησης απλοποιούνται αφού $\lambda_A = \lambda_B$ και $\mu_A = \mu_B$ για κάθε σύστημα. Επιλύουμε το πρόβλημα ελαχιστοποίησης του μέσου χρόνου απόκρισης δοσοληψιών των τριών συστημάτων. Αποδεικνύεται ότι $T_2 < T_3 < T_1$. Η διαπίστωση αυτή είναι σύμφωνη με τα αποτελέσματα των προσομοιώσεων.

3.6 Επισκόπηση Σχετικής Βιβλιογραφίας

Εχουν προταθεί πολλοί οικονομικοί αλγόριθμοι για κατανομή φόρτου (economic load balancing algorithms) σε υπολογιστικά συστήματα. Οι Waldspurger *et. al.* σχεδίασαν το Spawn [26] μία από τις πρώτες υπολογιστικές οικονομίες. Το Spawn είναι ένα σύστημα κατανομής φόρτου (load sharing system) για τοπικά δίκτυα (local area networks). Το σύστημα χρησιμοποιεί τον μηχανισμό πλειστηριασμού για να κατανέμει υπολογιστικό χρόνο σε ένα σύνολο από δοσοληψίες $T_1, T_2, ...T_N$. Στο Spawn δεν κοστολογείται το εύρος επικοινωνίας (communication bandwidth) κάτι που θα ήταν απαραίτητο σε μεγαλύτερα δυναμικά ετερογενή συστήματα.

Οι Malone *et. al.* σχεδίασαν το σύστημα Enterprice [16]στο οποίο οι κόμβοι συνεργάζονται για την κατανομή υπολογιστικού χρόνου. Το σύστημα Enterprice είναι μία οικονομία ανταλλαγής προϊόντων στην οποία δεν έχει εισαχθεί η έννοια του χρήματος.

Το Enterprice όπως και το Spawn δε χειρίζεται το πρόβλημα κατανομής επικοινωνιακών πόρων (communication resourses).

Οικονομία κατανομής φόρτου έχει σχεδιαστεί και από τους Ferguson *et. al.* [4, 5, 6, 7]. Στην οικονομία αυτή υπάρχουν Ν επεξεργαστές (προμηθευτές) οι οποίοι πωλούν υπολογιστικό χρόνο (CPU time) και εύρος επικοινωνίας (communication bandwidth) σε ένα σύνολο από δοσοληψίες (καταναλωτές). Ο μηχανισμός κοστολόγησης στηρίζεται στη διεξαγωγή πλειστηριασμών από τους επεξεργαστές. Στους πλειστηριασμούς πλειοδοτούν οι δοσοληψίες.

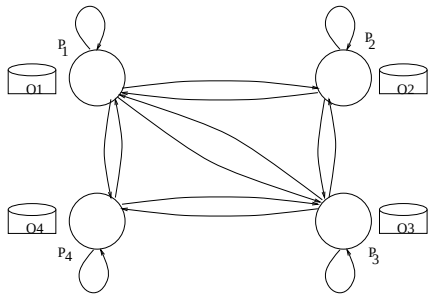
Το Lottery Scheduling και το Stride Scheduling [27, 28] είναι μηχανισμοί ανάθεσης υπολογιστικού χρόνου σε δοσοληψίες που στηρίζονται στην κοστολόγηση του υπολογιστικού χρόνου και μπορούν να γενικευθούν ώστε να διαχειρίζονται ένα σύνολο από διαφορετικούς πόρους.

Πρόσφατα οι Stonebraker *et. al.* μελέτησαν το πρόβλημα της εκτέλεσης επερωτήσεων σε μεγάλα κατανεμημένα συστήματα που αποτελούνται από χιλιάδες κόμβους. Εχουν επιλέξει ένα οικονομικό παράδειγμα παρόμοιο με αυτό που περιγράφηκε στις εργασίες [6, 7] για την σχεδίαση μιας κατανεμημένης βάσης δεδομένων, που ονομάζουν Mariposa [23]. Εχουν σχεδιάσει επαρκής μηχανισμούς πλειστηριασμού για την προσπέλαση των δεδομένων της βάσης και την αγορά υπολογιστικού χρόνου από τους πελάτες (χρήστες) του συστήματος

Κεφάλαιο 4

Οικονομία Διαχείρισης Δεδομένων

4.1 Περιγραφή Προβλήματος



Σχήμα 4.1: Μοντέλο Προβλήματος

Ας θεωρήσουμε κατανεμημένο σύστημα επεξεργασίας δοσοληψιών που αποτελείται από 4 κόμβους P_1, P_2, P_3, P_4. Κάθε κόμβος διαθέτει έναν επεξεργαστή και μία πηγή φόρτου. Η υπολογιστική δύναμη του επεξεργαστή P_i καθορίζεται από την παράμετρο $r_i, i \in [1..4]$. Οι κόμβοι συνδέονται με ένα ‘‘point-to-point δίκτυο το οποίο καθορίζεται από το σύνολο ακμών $E = e_{ij}$ όπου $i, j \in [1..4]$ και $i \neq j$. Η ακμή e_{ij} παριστά το μονόδρομο σύνδεσμο επικοινωνίας μεταξύ των επεξεργαστών P_i, P_j. Ο σύνδεσμος αυτός παρουσιάζει καθυστέρηση d_{ij} seconds per byte. Η βάση δεδομένων αποτελείται από 4 σχέσεις O_1, O_2, O_3, O_4 οι οποίες έχουν ανατεθεί στους κόμβους $P_1, P2, P3, P4$ αντίστοιχα. Εχουν οριστεί 2 κλάσεις δοσοληψιών C_A, CB. Οι δοσοληψίες της κλάσης C_A προσπελαύνουν κάποιο από τα δεδομένα O_1, O_2 ενώ οι δοσοληψίες της κλάσης C_B προσπελαύνουν κάποιο

από τα δεδομένα O_1, O_3, O_4. Θεωρούμε οτι οι υπολογιστικές απαιτήσεις των δοσοληψιών είναι γνωστές κατά την είσοδό τους στο σύστημα. Ο υπολογιστικός χρόνος που απαιτείται για την εξυπηρέτηση δοσοληψίας της κλάσης C_A σε ιδανικό επεξεργαστή P_i ($r_i = 1$) είναι μ_A ενώ αντίστοιχα μ_B είναι ο χρόνος που απαιτείται για την εξυπηρέτηση δοσοληψίας της κλάσης C_B. Ο ρυθμός άφιξης δοσοληψιών στους κόμβους του συστήματος παρουσιάζει μεταβολές στο χρόνο και είναι διαφορετικός για κάθε κλάση του συστήματος. Θεωρούμε οτι ο ρυθμός άφιξης δοσοληψιών της κλάσης C_A είναι κάθε χρονική στιγμή ίδιος για όλους τους κόμβους του συστήματος. Ανάλογα ο ρυθμός άφιξης δοσοληψιών της κλάσης C_B είναι κάθε χρονική στιγμή ίδιος για όλους τους κόμβους του συστήματος. Το πρόβλημα της διαχείρισης δεδομένων σε αυτό το κατανεμημένο σύστημα ορίζεται ως η ελαχιστοποίηση του μέσου χρόνου απόκρισης των κλάσεων δοσοληψιών C_A, C_B έχοντας τις ακόλουθες παραμέτρους ελέγχου :

- αριθμός αντιγράφων κάθε σχέσης.
- ανάθεση των αντιγράφων στους κόμβους
- δρομολόγηση των δοσοληψιών

Ο χρόνος απόκρισης μιας δοσοληψίας (της κλάσης C_j) στον επεξεργαστή P_i ορίζεται ως το άθροισμα του χρόνου αναμονής της δοσοληψίας (waiting time) στον επεξεργαστή P_i και του χρόνου εξυπηρέτησης ($\frac{\mu_j}{r_i}$) της δοσοληψίας (service time) στον επεξεργαστή. Δεδομένα του προβλήματος θεωρούνται οι ρυθμοί άφιξης λ_{A_j} των δοσοληψιών των κλάσεων C_A, C_B αντίστοιχα τη χρονική στιγμή t_j. Γνωρίζουμε επίσης και τους χρόνους εξυπηρέτησης $\frac{\mu_A}{r_i}$, $\frac{\mu_B}{r_i}$ των δοσοληψιών των κλάσεων C_A, C_B αντίστοιχα στον επεξεργαστή P_i Το πρόβλημα της διαχείρισης των δεδομένων του συστήματος θα επιλυθεί με τη σχεδίαση ενός δυναμικού αλγορίθμου αντιγραφής των σχέσεων της βάσης και δρομολόγησης των εργασιών στους κόμβους του συστήματος (dynamic replication and routing algorithm). Ο αλγόριθμος αυτός είναι ευαίσθητος στις μεταβολές του φόρτου του συστήματος και βασιζόμενος κυρίως σε αυτές θα παίρνει αποφάσεις για την αντιγραφή των σχέσεων της βάσης δεδομένων και για την δρομολόγηση των δοσοληψιών ώστε να επιταχύνεται ο χρόνος εκτέλεσης των εργασιών του συστήματος. Ο αλγόριθμος θα πρέπει να λαμβάνει επίσης υπόψη, τον υπάρχοντα φόρτο των επεξεργαστών, τον αναμενόμενο χρόνο εξυπηρέτησης της δρομολογούμενης δοσοληψίας σε κάθε επεξεργαστή, καθώς και την καθυστέρηση για την μεταβίβαση της δοσοληψίας σε κάθε επεξεργαστή. Θα πρέπει επίσης να γνωρίζει το φόρτο

που έχει συσσωρευτεί σε κάθε κόμβο για κάθε μία από τις σχέσεις που διαθέτει ο κόμβος καθώς και τη διαφορά φόρτου που παρουσιάζουν οι κλάσεις δοσοληψιών.

Οι περισσότερες από τις αναφερθείσες παραμέτρους δεν είναι σταθερές στο χρόνο οπότε ο αλγόριθμος θα πρέπει να αντιλαμβάνεται τις μεταβολές τους.

Η σχεδίαση ενός οικονομικού αλγορίθμου διαχείρισης δεδομένων και δρομολόγησης δοσοληψιών μειώνει την πολυπλοκότητα του προβλήματος αφού η διαδικασία λήψης αποφάσεων λαμβάνει χώρα σε ένα δυναμικό μη κεντρικοποιημένο περιβάλλον από αγοραστές και πωλητές, όπου οι τιμές των αγαθών διαμορφώνονται ανάλογα με τις παραμέτρους απόδοσης του συστήματος που έχουν περιγραφεί παραπάνω.

4.2 Οικονομικός αλγόριθμος αντιγραφής δεδομένων και δρομολόγησης δοσοληψιών DRR

Οπως μια ανθρώπινη οικονομία, το υπολογιστικό σύστημα διαθέτει ένα σύνολο διακριτών αγαθών, τα οποία πρέπει να διατεθούν σε ένα σύνολο καταναλωτών. Τα αγαθά του υπολογιστικού συστήματος είναι τα δεδομένα (data), ο υπολογιστικός χρόνος (CPU time) και το εύρος επικοινωνίας (communication bandwidth). Οι κόμβοι του συστήματος είναι οι πωλητές της οικονομίας ενώ οι κλάσεις δοσοληψιών θεωρούνται αγοραστές. Στόχος των κλάσεων δοσοληψιών είναι η εξυπηρέτησή τους με το μικρότερο δυνατό κόστος ενώ στόχος των κόμβων είναι η μεγιστοποίηση των κερδών τους. Οι κόμβοι κερδίζουν πουλώντας δεδομένα και υπολογιστικό χρόνο σε δοσοληψίες που εκτελούνται τοπικά. Χρησιμοποιούν τα κέρδη που αποκομίζουν, για την ενοικίαση ‘‘μη τοπικών’’ δεδομένων όταν θεωρούν οτι ανάλογη επένδυση, μακροπρόθεσμα, θα αυξήσει το εισόδημα τους. Η ενοικίαση ενός αντιγράφου γίνεται μετά από συμφωνία με τον ιδιοκτήτη (owner) του δεδομένου ή τον κάτοχο ενός άλλου αντιγράφου. Η τιμή ενοικίασης πρέπει να βρίσκει σύμφωνα και τα δύο μέλη του ‘‘συμβολαίου’’ και ανανεώνεται με την πάροδο του χρόνου ώστε να ανταποκρίνεται στην αξία του αγαθού. Οι κόμβοι προσφέρουν για ενοικίαση κάποιο από τα τοπικά τους δεδομένα ή αντίγραφο δεδομένου που έχουν ενοικιάσει, όταν η τιμή ενοικίασης εγγυάται αύξηση των κερδών τους. Σε κάθε δοσοληψία της κλάσης C_j που εισέρχεται στο σύστημα διατίθεται κεφάλαιο B_j. Η δοσοληψία εκμεταλλεύεται το χρηματικό αυτό ποσό για την αγορά δεδομένων, υπολογιστικού χρόνου και εύρους επικοινωνίας.Στόχος της δοσοληψίας είναι η ελαχιστοποίηση του

κόστους αγοράς αγαθών, οπότε επιλέγει να εξυπηρετηθεί από τον κόμβο που καλύπτει τις απαιτήσεις της και προσφέρει την μικρότερη δυνατή τιμή. Η τιμή κάθε κόμβου αποτελεί συνάρτηση του φόρτου του κόμβου, της ζήτησης των δεδομένων που προσφέρει και του κόστους μεταβίβασης της δρομολογούμενης δοσοληψίας. Οι κλάσεις δοσοληψιών έχουν τις ίδιες προτιμήσεις σε αγαθά (utility function) αλλά μπορεί να διαθέτουν διαφορετικά κεφάλαια. Το κεφάλαιο κάθε κλάσης καθορίζεται από τον διαχειριστή του συστήματος και αντικατοπτρίζει τις υπολογιστικές απαιτήσεις κάθε κλάσης. Η δοσοληψία αποτυγχάνει να εξυπηρετηθεί όταν δεν διαθέτει ισχυρό κεφάλαιο (όλοι οι κόμβοι ζητούν περισσότερα από όσα μπορεί να πληρώσει). Το κεφάλαιο που διαθέτει κάθε κλάση αποτελεί δείκτη της προτεραιότητας των δοσοληψιών. Οσο μεγαλύτερο είναι το κεφάλαιο μιας κλάσης τόσα περισσότερα αγαθά ή ποσοστό ενός αγαθού δύναται να αγοράσει η κλάση αυτή και τόσο μικρότερη είναι η πιθανότητα απόρριψης δοσοληψιών της κλάσης. Η οικονομική πολιτική που ακολουθείται για την διαμόρφωση των τιμών είναι κοινή για όλους τους κόμβους του συστήματος. Κανένας κόμβος δεν υιοθετεί δόλια συμπεριφορά, δεν χρεώνει δηλαδή τα αγαθά του σε τιμές δυσανάλογες της αξίας τους.

Περιγραφή οικονομικού αλγορίθμου DRR

Θεωρώ σύστημα N επεξεργαστών $P_1, P_2, .., P_N$ όπου r_i η υπολογιστική δύναμη του επεξεργαστή $P_i, i \in [1..N]$. Η βάση δεδομένων αποτελείται από M σχέσεις $O_1, O_2, .., O_M$. $S(O_i)$ είναι το πλήθος σελίδων της σχέσης O_i ενώ $P_{ij}(O_i)$ είναι ο χρόνος προσπέλασης της σχέσης O_i στο χρονικό διάστημα t_j στον κόμβο P_i. e_{ij} είναι ο σύνδεσμος επικοινωνίας μεταξύ των επεξεργαστών P_i, P_j. d_{ij} είναι η καθυστέρηση του συνδέσμου επικοινωνίας e_{ij} η οποία δεν μεταβάλλεται στο χρόνο.

Εχουν οριστεί Λ κλάσεις δοσοληψιών $C_A, C_B, C_\Gamma, .., C_\Lambda$. μ_i είναι ο αναμενόμενος χρόνος εξυπηρέτησης δοσοληψίας της κλάσης C_i σε ιδανικό επεξεργαστή. Η συχνότητα άφιξης δοσοληψιών της κλάσης C_i το χρονικό διάστημα t_j είναι και είναι λ_{ij} και είναι ίδια για όλους τους κόμβους (πηγές φόρτου) του συστήματος. Οταν μια δοσοληψία της κλάσης $C_j, j \in [1..\Lambda]$ εισέρχεται στο σύστημα κατά το χρονικό διάστημα t_α, διαθέτει κεφάλαιο B_j το οποίο εκμεταλλεύεται για να αγοράσει, με το μικρότερο εφικτό κόστος, δεδομένα, υπολογιστικό χρόνο και εύρος επικοινωνίας. Η δοσοληψία θα επιλέξει να δρομολογηθεί στον κόμβο που μπορεί να καλύψει τις απαιτήσεις της και προσφέρει την μικρότερη τιμή για τις υπηρεσίες του. Η συνάρτηση χρησιμότητας (utility function) αντανακλά τις προτιμήσεις της κλάσης C_j και καθορίζει τις αποφάσεις δρομολόγησης:

Utility function =

$$\min_{\forall i}\{C_i\} = \min_{\forall i}\{J*(WT_{ai}+ST_i+d_{ki}+\sum P_{ai}(O_p))\} = \min_{\forall i}\{J*(WT_{ai}+\frac{\mu_j}{r_i}+d_{ki}+\sum P_{ai}(O_p))\}$$
$$i \in [1, N]$$

όπου WT_{ai} είναι ο χρόνος αναμονής (waiting time) της δρομολογούμενης δοσοληψίας στον κόμβο P_i το χρονικό διάστημα t_α. Ο χρόνος αναμονής ορίζεται ως το άθροισμα των χρόνων εξυπηρέτησης των δοσοληψιών που έχουν ήδη δρομολογηθεί στον επεξεργαστή και περιμένουν να εκτελεστούν. Η παράμετρος $ST_i = \frac{\mu_j}{r_i}$ δηλώνει τον αναμενόμενο χρόνο εξυπηρέτησης (service time) της δοσοληψίας της κλάσης C_j στον κόμβο P_i. Η παράμετρος d_{ki} είναι ο χρόνος μεταβίβασης της δοσοληψίας από την πηγή φόρτου P_k στον κόμβο P_i. Το κόστος αυτό είναι μηδενικό αν η δοσοληψία εκτελεστεί τοπικά. Το άθροισμα $\sum P_{ai}(O_p)$ δηλώνει το συνολικό κόστος των σχέσεων που επιθυμεί να προσπελάσει η δρομολογούμενη δοσοληψία. $P_{ai}(\mathrm{O}_p)$ είναι ο χρόνος που απαιτείται για προσπέλαση της σχέσης O_p στο χρονικό διάστημα t_α στον κόμβο P_i. Ο πολλαπλασιαστής J έχει τιμή 1 δραχμή ανά δεπτερόλεπτο.

Αν η δοσοληψία δεν διαθέτει αρκετά μεγάλο κεφάλαιο ώστε να μπορεί να πληρώσει κάποιον από τους κόμβους, απορρίπτεται από το σύστημα.

Στην αρχή κάθε χρονικού διαστήματος δίνεται η δυνατότητα στους κόμβους του συστήματος να ενοικιάσουν δεδομένα τα οποία δεν διαθέτουν αλλά παρουσίασαν αυξημένη ζήτηση στο προηγούμενο χρονικό διάστημα.

Τα δεδομένα αυτά θα ζητηθούν για ενοικίαση από τους ιδιοκτήτες κόμβους (κόμβοι στους οποίους έχει ανατεθεί το δεδομένο από τον διαχειριστή του συστήματος) ή από κόμβους που κατέχουν αντίγραφο, το οποίο ενοικίασαν σε προηγούμενο χρονικό διάστημα. Ο κόμβος που επιθυμεί να ενοικιάσει αντίγραφο μιας σχέσης O_p, θα συνάψει ‘‘συμβόλαιο’’ με τον κάτοχο της σχέσεις που ζητεί μικρότερο ενοίκιο. Η επιλογή αυτή γίνεται για να μειωθεί το κόστος ενοικίασης.

Εστω P_j ο κόμβος που επιθυμεί να ενοικιάσει αντίγραφο της σχέσης O_p το χρονικό διάστημα t_α και P_i ο πιο ‘‘φθηνός’’ κάτοχος της σχέσης το χρονικό αυτό διάστημα. Το ‘‘ενοίκιο’’ (LeasePrice) που ζητάει ο κόμβος P_i δίνεται από τον ακόλουθο τύπο:

$$LeasePrice_{ai} = P_{ai}(O_p) * Service_{(\alpha-1)i}(O_p) + S(O_p) * d_{ij}$$

όπου $P_{\alpha i}(\mathrm{O}_p)$ η τιμή της σχέσης O_p στον κόμβο P_i το χρονικό διάστημα t_α, $Service_{(\alpha-1)i}(\mathrm{O}_p)$ ο αριθμός δοσοληψιών που εκτελέστηκαν στον κόμβο P_i το προηγούμενο

χρονικό διάστημα $t_{\alpha-1}$ και ζητούσαν την σχέση O_p και $S(O_p) * d_{ij}$ το κόστος μεταβίβασης των σελίδων της σχέσης O_p από τον κόμβο P_i στον υποψήφιο ενοικιαστή P_j. Το ενοίκιο ισούται στην χειρότερη περίπτωση με το ποσό που θα χάσει ο κόμβος P_i αν αριθμός δοσοληψιών ίσος με τον αριθμό των δοσοληψιών που εκτελέστηκαν τοπικά το χρονικό διάστημα $t_{\alpha-1}$ και ζητούσαν την σχέση O_p δρομολογηθούν στο χρονικό διάστημα t_α στον κόμβο P_j (συν το κόστος μεταβίβασης της σχέσης).

Το προηγούμενο χρονικό διάστημα ο κόμβος P_j δέχτηκε πλήθος δοσοληψιών ίσο με $Demand_{(\alpha-1)j}(O_p)$ οι οποίες επιθυμούσαν να προσπελάσουν τη σχέση O_p. Οι δοσοληψίες αυτές δεν μπόρεσαν να εκτελεστούν τοπικά (λόγω απουσίας της σχέσης O_p από τον κόμβο P_j) και έτσι δρομολογήθηκαν σε άλλους κόμβους. Ο κόμβος P_j ζημειώθηκε κατά:

$$P_{ai}(O_p) * Demand_{(a-1)j}(O_p) + d_{ij} * Demand_{(a-1)j}(O_p)\text{δρχ}.$$

όπου $P_{ai}(O_p) * Demand_{(\alpha-1)j}(O_p)$ το ποσό που δαπάνησαν $Demand_{(\alpha-1)j}(O_p)$ δοσοληψίες για την αγορά του αγαθού O_p σε κάποιον άλλο κόμβο (κάτοχο της σχέσης O_p) και $d_{ji} * Demand_{(\alpha-1)j}(O_p)$ το ποσό που δαπάνησαν οι δοσοληψίες για την μεταβίβασή τους από τον κόμβο P_j σε κάποιον άλλο κόμβο. Θεωρούμε οτι οι δοσοληψίες δρομολογήθηκαν στον κόμβο P_i. Αν οι δοσοληψίες εκτελούνταν τοπικά, ανάλογο ποσό χρημάτων θα μπορούσε να είναι η αμοιβή του κόμβου P_j. Αρα ο κόμβος P_j δύναται να πληρώσει ενοίκιο (OfferPrice) ίσο με:

$$OfferPrice_{\alpha j} = P_{ai}(O_p) * Demand_{(a-1)j}(O_p) + d_{ji} * Demand_{(a-1)j}(O_p)$$

Αν $LeasePrice_{\alpha i} < OfferPrice_{\alpha j}$ συνάπτεται συμβόλαιο μεταξύ των κόμβων P_i, P_j και ο κόμβος P_j αποκτά αντίγραφο της σχέσης O_p για το χρονικό διάστημα t_α . Ο κόμβος P_j διαθέτει τη σχέση O_p σε τιμή ίση με :

$$P_{aj}(O_p) = P_{ai}(O_p) * \frac{r_i}{r_j}$$

Παρατηρούμε οτι αν ο κόμβος P_j έχει υπολογιστική δύναμη μεγαλύτερη από τον κόμβο P_i $(r_j > r_i)$ θα προσφέρει τη σχέση O_p σε χαμηλότερη τιμή. Αν ο κόμβος P_j έχει μεγαλύτερη υπολογιστική δύναμη και εγγυάται μικρότερη τιμή για την σχέση O_p από τον κόμβο P_i, αναμένεται να προσελκύσει μεγαλύτερο αριθμό δοσοληψιών που ζητούν να προσπελάσουν την σχέση O_p. Ετσι θα επιτύχει να αποκομίσει ποσό χρημάτων μεγαλύτερου του ενοικίου που κατέβαλε στον κόμβο P_i. Αν αντίθετα έχει μικρότερη υπολογιστική δύναμη και

ακριβότερη τιμή για την σχέση O_p από τον κόμβο P_i θα προσελκύσει δοσοληψίες και θα αποσβέσει το ενοίκιο μόνο αν ο φόρτος του κόμβου P_i γίνει αρκετά μεγάλος ώστε να καλύψει τα προηγούμενα πλεονεκτήματά του (χαμηλή τιμή, μεγάλη υπολογιστική δύναμη).

Αν ο κόμβος P_j επιθυμεί να διατηρήσει αντίγραφο της σχέσης O_p μετά το τέλος του χρονικού διαστήματος t_α πρέπει να διαπραγματευτεί ξανά με τον κόμβο P_i (κόμβο από τον οποίο έχει ενοικιάσει το αντίγραφο) και να πληρώσει μέρος του ενοικίου για το χρονικό διάστημα $t_{\alpha+1}$.

Ο κόμβος P_i για το χρονικό διάστημα $t_{\alpha+1}$ απαιτεί ποσό ίσο με:

$$RenewalPrice_{(a+1)i} = P_{(a+1)i}(O_p) * Service_{ai}(O_p)$$

Το ποσό αυτό ισούται με το ποσό χρημάτων που θα χάσει στην χειρότερη περίπτωση ο κόμβος P_i αν σύνολο δοσοληψιών ίσο με τον αριθμό των δοσοληψιών που ζήτησαν την σχέση O_p το προηγούμενο χρονικό διάστημα και εξυπηρετήθηκαν στον κόμβο P_i δρομολογηθούν στο διάστημα $t_{\alpha+1}$ στον κόμβο P_j. Ο κόμβος P_j δύναται να πληρώσει, για την ανανέωση του ‘‘ενοικιοστασίου’’, ποσό ίσο με:

$$ReofferPrice_{(a+1)j} = P_{(a+1)j}(O_p) * Service_{aj}(O_p)$$

Το ποσό αυτό ισούται με το ποσό των χρημάτων που θα αποκομίσει ο κόμβος P_j το χρονικό διάστημα $t_{\alpha+1}$ αν εξυπηρετήσει σύνολο δοσοληψιών ίσο με τον αριθμό των δοσοληψιών που ζήτησαν την σχέση O_p και εκτελέστηκαν τοπικά το προηγούμενο χρονικό διάστημα. Αν

$$RenewalPrice_{(\alpha+1)i} \leq ReofferPrice_{(\alpha+1)j}$$

τότε το συμβόλαιο των κόμβων P_i, P_j ανανεώνεται για το χρονικό διάστημα $t_{\alpha+1}$ και ο κόμβος P_j διατηρεί το αντίγραφο της σχέσης O_p. Ανάλογη διαδικασία ανανέωσης του συμβολαίου πρέπει να ακολουθηθεί και στα επόμενα χρονικά διαστήματα $t_{\alpha+2}, t_{\alpha+3}$ κ.ο.κ.

Αν

$$RenewalPrice_{\alpha(i+1)} > ReofferPrice_{(\alpha+1)j}$$

Ο κόμβος P_j δε θεωρεί πλέον προσοδοφόρο να διατηρεί αντίγραφο της σχέσης O_p. Ο κόμβος P_j αναγκάζεται να σβήσει το αντίγραφο της σχέσης O_p από τον τοπικό δίσκο και να ενημερώσει τους υπόλοιπους κόμβους οτι δεν εξυπηρετεί πλέον δοσοληψίες που επιθυμούν να προσπελάσουν τη σχέση O_p.

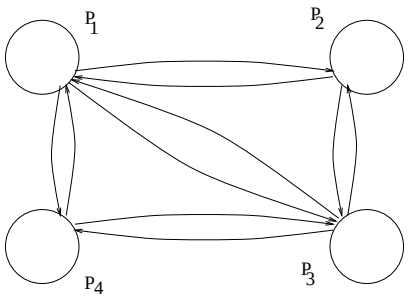
Στον αλγόριθμο DRR ο χρόνος αναμονής της δοσοληψίας j στον επεξεργαστή P_i αποτελεί παράμετρο στη λήψη της απόφασης δρομολόγησης. Η παράμετρος αυτή αντικατοπτρίζει την ζήτηση που παρατηρείται στον κόμβο P_i και αποτελεί μέτρο του υπολογιστικού φόρτου που έχει ήδη δρομολογηθεί στον κόμβο αυτό. Οι τιμές στις οποίες οι κόμβοι προσφέρουν τα αγαθά τους διαμορφώνονται με τρόπο ώστε να αντικατοπτρίζουν τις δυνατότητες προσφοράς των κόμβων. Για παράδειγμα η τιμή των αντιγράφων σχέσεων ενός κόμβου αποτελεί συνάρτηση της υπολογιστικής δύναμης του κόμβου αυτού. Επίσης στην κοστολόγηση των δοσοληψιών του, ένας κόμβος λαμβάνει υπόψη του τόσο το χρόνο αναμονής όσο και το χρόνο εξυπηρέτησης της δρομολογούμενης δοσοληψίας. Οι χρόνοι αυτοί αποτελούν συνάρτηση του πλήθους των δοσοληψιών που έχουν δρομολογηθεί στον κόμβο, της υπολογιστικής δύναμης του κόμβου, καθώς και των δεδομένων που διαθέτει (εκτελεί δοσοληψίες που προσπελαύνουν δεδομένα των οποίων έχει αντίγραφα). Η απόφαση για την ενοικίαση ενός αντιγράφου βασίζεται κι αυτή στη ζήτηση και στην προσφορά που παρατηρείται στους κόμβους. Ο κάτοχος πρέπει να έχει συσσωρευμένο σχετικά μεγάλο αριθμό δοσοληψιών που ζητούν τη σχέση ώστε να μη ζημιωθεί από τον ανταγωνισμό που θα προκαλέσει ένα καινούριο αντίγραφο. Ανάλογα ο πιθανός ενοικιαστής πρέπει να έχει μεταβιβάσει σε άλλους κόμβους αρκετές δοσοληψίες που επιθυμούν μια συγκεκριμένη σχέση, πριν αποφασίσει οτι είναι προσοδοφόρο να ενοικιάσει αντίγραφο της σχέσης αυτής.

Οι κόμβοι είναι αξιόπιστοι και εγγυώνται οτι δεν πρόκειται να απαιτήσουν από κάποια δοσοληψία ποσό μεγαλύτερο της τιμής που ζήτησαν για την εκτέλεσή της πριν ληφθεί η απόφαση δρομολόγησης (price guarantees). Το σύστημα οδηγείται σε κατάσταση ισορροπίας όπου η ζήτηση σε ένα επεξεργαστή ισούται με την προσφορά.

4.3 Μοντέλα προσομοίωσης

Ο οικονομικός αλγόριθμος DRR αποτέλεσε το δυναμικό αλγόριθμο διαχείρισης δεδομένων και δρομολόγησης δοσοληψιών σε πειράματα προσομοίωσης που διεξήχθηκαν χρησιμοποιώντας το σύστημα επεξεργασίας δοσοληψιών του Σχήματος 4.2.

Το σύστημα επεξεργασίας δοσοληψιών αποτελείται από 4 κόμβους P_1, P_2, P_3, P_4. Ολοι οι κόμβοι λειτουργούν ως πηγές φόρτου ενώ παράλληλα εκτελούν τις δοσοληψίες που υποβάλλονται στο καθένα από 40 τερματικά. Οι κόμβοι έχουν από ένα επεξεργαστή



Σχήμα 4.2: Μοντέλο Προσομοίωσης

που εφαρμόζει FCFS sceduling policy. Ο μέγιστος βαθμός πολυπρογραμματισμού είναι 200. Η βάση δεδομένων αποτελείται από 4 σχέσεις Α,Β,Γ,Δ. Στους κόμβους, τα δεδομένα κατανέμονται με την μέθοδο round-robin σε δύο δίσκους, ενώ το ημερολόγιο (log) διατηρείται σε ξεχωριστό δίσκο. Οι δίσκοι δίνουν μέσο χρόνο προσπέλασης ίσο με 10msec. Κάθε κόμβος διαθέτει ενταμιευτή δεδομένων μεγέθους 400,000 σελίδων. Το κόστος επικοινωνίας του κόμβου P_1 με τους κόμβους P_2, P_3, P_4 είναι 2.0, 2.5, 3.0 αντίστοιχα. Το κόστος επικοινωνίας του κόμβου P_2 με τους κόμβους P_1, P_3, P_4 είναι 2.5, 3.0 και 2.5 αντίστοιχα. Το κόστος επικοινωνίας του κόμβου P_3 με τους κόμβους P_1, P_2, P_4 είναι 2.0, 2.5 και 2.0 αντίστοιχα ενώ το κόστος επικοινωνίας του κόμβου P_4 με τους κόμβους P_1, P_2, P_3 είναι είναι 2.0, 2.5, 3.0 αντίστοιχα. Εχουν οριστεί 4 κλάσεις δεδομένων C_A, C_B, C_Γ, C_D οι οποίες επιθυμούν να προσπελάσουν τις σχέσεις Α, Β, Γ, Δ αντίστοιχα. Οι δοσοληψίες όλων των κλάσεων εκτελούν κατά μέσο όρο 10,000,000 εντολές και από 1 προσπέλαση/ανάγνωση. Ο χρόνος προσομοίωσης έχει διαιρεθεί σε διαστήματα στα οποία ο ρυθμός άφιξης των δοσοληψιών μεταβάλλεται. Οι κλάσεις δοσοληψιών C_i παρουσιάζουν σε κάθε χρονικό διάστημα την ίδια συχνότητα άφιξης σε όλους τους κόμβους του συστήματος. Οι κλάσεις δοσοληψιών έχουν το ίδιο κεφάλαιο αλλά είναι δυνατό να παρουσιάζουν διαφορές στον ρυθμό άφιξης και στις αρχικές τιμές των δεδομένων τους.

4.4 Αποτελέσματα

Στις προσομοιώσεις που μελετώνται παρακάτω οι σχέσεις Α, Β, Γ, Δ έχουν ανατεθεί στους κόμβους P_1, P_2, P_3, P_4 αντίστοιχα. Αρχικά διεξήχθηκαν πειράματα για να μελετηθεί η επίδραση των τιμών των δεδομένων και του φόρτου του συστήματος στη λήψη αποφάσεων

διαχείρισης δεδομένων και δρομολόγησης δοσοληψιών.

	$RespC_A$	$RespC_{\text{B}}$	$RespC_{\Gamma}$	$RespC_{\Delta}$
Πείραμα 1	0.86	0.99	0.77	0.63
Πείραμα 2	0.63	0.86	0.52	0.93
Πείραμα 3	0.71	0.96	0.73	0.71

Πίνακας 1α

	Α	Β	Γ	Δ
Πείραμα 1	P_1, P_2, P_3	P_2, P_4	P_2, P_3, P_4	P_1, P_2, P_3, P_4
Πείραμα 2	P_1, P_2, P_3	P_1, P_2	P_1, P_2, P_3, P_4	P_3, P_4
Πείραμα 3	P_1, P_2, P_3	P_2, P_4	P_1, P_2, P_3, P_4	P_1, P_2, P_3, P_4

Πίνακας 1β

Στον Πίνακα 1α παρουσιάζονται τα αποτελέσματα 3 πειραμάτων προσομοίωσης. Και στα τρία πειράματα η μεταβολή του ρυθμού άφιξης δοσοληψιών στη διάρκεια του χρόνου προσομοίωσης είναι ίδια για όλες τις κλάσεις και όλους τους κόμβους του συστήματος. Συγκεκριμένα για κάθε πείραμα ο ρυθμός άφιξης όλων των κλάσεων δοσοληψιών αυξάνεται ως εξής:

Ρυθμός Άφιξης : $1.0[0, 500]$

$\frac{10}{9}[500, 1000]$

$\frac{10}{8}[1000, 2000]$

$\frac{10}{7}[2000, 2500]$

$\frac{10}{6}[2500, 3600]$

Στο πρώτο πείραμα η τιμή των σχέσεων Α, Β, Γ, Δ που έχουν ανατεθεί στους κόμβους P_1, P_2, P_3, P_4 αντίστοιχα είναι 10 δρχ. Η σχέση Α κατά τη διάρκεια της προσομοίωσης επιτυγχάνει τη δημιουργία αντιγράφων της στους κόμβους P_2, P_3, η σχέση Β στον κόμβο P_4, η σχέση Γ στους κόμβους P_2, P_4 ενώ η Δ επιτυγχάνει να αντιγραφεί σε όλους τους κόμβους (Πίνακας 1β). Μικρότερο μέσο χρόνο απόκρισης στο πείραμα αυτό παρουσιάζει η κλάση C_{Δ} η οποία ζητεί να προσπελάσει την σχέση Δ. Η σχέση Δ έχει αντιγραφεί σε περισσότερους κόμβους.

Στο δεύτερο πείραμα η τιμή της σχέσης Γ στον κόμβο P_3 μειώνεται σε 5 δρχ. Κατά την διάρκεια της προσομοίωσης η σχέση Α επιτυγχάνει να αντιγράφει στους κόμβους P_2, P_3, η

σχέση Β στον κόμβο P_1, η σχέση Γ επιτυγχάνει να αντιγράφει σε όλους τους κόμβους ενώ η σχέση Δ στον κόμβο P_3. Στον Πίνακα 1α παρατηρούμε ότι ο μέσος χρόνος απόκρισης της κλάσης δοσοληψιών C_Γ μειώνεται από 0.77 seconds (πρώτο πείραμα) σε 0.52 seconds και αποτελεί το μικρότερο χρόνο που επιτυγχάνεται για κλάση του συστήματος. Το γεγονός αυτό οφείλεται στην μείωση της τιμής της σχέσης Γ που επέτρεψε την αντιγραφή της σχέσης σε όλους τους κόμβους. Η μείωση της τιμής της σχέσης Γ ζημίωσε ιδιαίτερα την κλάση δοσοληψιών C_Δ. Το γεγονός ότι η τιμή της σχέσης Δ είναι σχετικά ακριβότερη και όλοι οι κόμβοι προσφέρουν εξυπηρέτηση σε δοσοληψίες της κλάσης C_Γ, περιορίζει την αντιγραφή της σχέσης Δ στον κόμβο P_3 και έτσι επιβραδύνει την εκτέλεση των δοσοληψιών της κλάσης C_Δ.

Στο πείραμα 3 η τιμή της σχέσης Γ παραμένει 5 δρχ ενώ μειώνεται σε 5 δρχ και η τιμή της σχέσης Δ. Παρατηρούμε ότι οι σχέσεις Γ, Δ στο πείραμα αυτό επιτυγχάνουν να αντιγράφουν σε όλους τους κόμβους του συστήματος. Οι μέσοι χρόνοι απόκρισης δοσοληψιών των κλάσεων C_Γ, C_Δ είναι ανάλογοι 0.73 και 0.71 seconds αντίστοιχα. Παρατηρούμε ότι ενώ και στο δεύτερο και στο τρίτο πείραμα η σχέση Γ έχει αντιγραφεί σε όλους τους κόμβους του συστήματος (Πίνακας 1β) ο μέσος χρόνος απόκρισης της κλάσης δοσοληψιών C_Γ είναι μικρότερος στο τρίτο πείραμα. Αυτό οφείλεται στο ότι στο τρίτο πείραμα ο ανταγωνισμός της κλάσης C_Γ έχει αυξηθεί σε κάθε κόμβο (υπάρχει πλέον αντίγραφο της σχέσης Δ σε όλους τους κόμβους). Ανάλογα συμπεράσματα προκύπτουν από την μελέτη της συμπεριφοράς της κλάσης δοσοληψιών C_Δ στο πρώτο και στο τρίτο πείραμα.

	$RespC_A$	$RespC_\text{B}$	$RespC_\Gamma$	$RespC_\Delta$
Πείραμα 1	0.38	0.38	0.31	0.43
Πείραμα 2	0.45	0.36	0.39	0.32
Πείραμα 3	0.41	0.37	0.41	0.37

Πίνακας 2α

	A	B	Γ	Δ
Πείραμα 1	P_1, P_3, P_4	P_2, P_4, P_3	P_1, P_2, P_3	P_1, P_2, P_4
Πείραμα 2	P_1, P_4	P_1, P_2, P_3	P_1, P_3	P_1, P_2, P_3, P_4
Πείραμα 3	P_2, P_4	P_1, P_2, P_3, P_4	P_1, P_3	P_1, P_2, P_3, P_4

Πίνακας 2β

Στον πίνακα 2α παρουσιάζονται τα αποτελέσματα 3 πειραμάτων προσομοίωσης, στα οποία ο φόρτος του συστήματος είναι μικρότερος σε κάθε διάστημα του χρόνου προσομοίωσης συγκριτικά με τα πειράματα του Πίνακα 1α. Συγκεκριμένα ο ρυθμός άφιξης όλων των κλάσεων δοσοληψιών αυξάνεται ως εξής:

Ρυθμός Άφιξης : $\frac{10}{15}[0, 500]$
$\frac{10}{10}[500, 1000]$
$\frac{10}{9}[1000, 2000]$
$\frac{10}{8}[2000, 2500]$
$\frac{10}{7}[2500, 3600]$

Η μείωση του φόρτου του συστήματος έχει ως αποτέλεσμα οι μέσοι χρόνοι απόκρισης όλων των κλάσεων δοσοληψιών του Πίνακα 2α να είναι μικρότεροι συγκριτικά με τους αντίστοιχους χρόνους που παραθέτονται στον Πίνακα 1α. Τα πειράματα αυτά επιβεβαιώνουν για μικρότερο φόρτο του συστήματος τα συμπεράσματα που προέκυψαν από την μελέτη του Πίνακα 1α. Στο πρώτο πείραμα η τιμή όλων των σχέσεων είναι 10 δρχ. Κάθε σχέση επιτυγχάνει να αντιγραφεί σε δύο κόμβους (Πίνακας 2β). Η σχέση Α αντιγράφεται στους κόμβους P_3, P_4 η σχέση Β στους κόμβους P_3, P_4 η σχέση Γ στους κόμβους P_1, P_2 ενώ η σχέση Δ αντιγράφεται στους κόμβους P_1, P_2.

Στο δεύτερο πείραμα η τιμή της σχέσης Δ μειώνεται σε 5 δρχ με αποτέλεσμα η σχέση Δ να επιτύχει την αντιγραφή της σε όλους τους κόμβους και η κλάση δοσοληψιών C_Δ να μειώσει το μέσο χρόνο απόκρισης της από 0.42 σε 0.32 seconds. Οι σχέσεις Α, Γ στο δεύτερο πείραμα αντιγράφονται σε λιγότερους κόμβους με αποτέλεσμα οι μέσοι χρόνοι απόκρισης των κλάσεων δοσοληψιών C_A, C_Γ να αυξηθούν από 0.38, 0.31 seconds σε 0.54, 0.39 seconds αντίστοιχα.

Στο πείραμα 3 οι σχέσεις Β, Δ παρουσιάζουν την μικρότερη τιμή (5 δρχ) και αντίγραφα τους αναθέτονται σε όλους τους κόμβους. Οι μέσοι χρόνοι απόκρισης των κλάσεων C_B, C_Δ είναι ανάλογοι (0.37 seconds) και μικρότεροι από τους μέσους χρόνους απόκρισης των υπολοίπων κλάσεων δοσοληψιών.

Στα προηγούμενα πειράματα προσομοίωσης οι κλάσεις δοσοληψιών παρουσιάζουν την ίδια συχνότητα άφιξης στους κόμβους του συστήματος.

	$RespC_A$	$RespC_{\text{B}}$	$RespC_{\Gamma}$	$RespC_{\Delta}$
Πείραμα 1	0.35	0.40	0.97	1.11
Πείραμα 2	0.91	1.34	0.41	0.34

Πίνακας 3α

	Α	Β	Γ	Δ
Πείραμα 1	P_1, P_2, P_4	P_1, P_2	P_1, P_2, P_3, P_4	P_1, P_3, P_4
Πείραμα 2	P_1, P_2, P_3, P_4	P_2, P_3	P_3, P_4	P_1, P_2, P_3

Πίνακας 3β

Στα πειράματα του Πίνακα 3α η συχνότητα άφιξης των κλάσεων δοσοληψιών $C_{\text{A}}, C_{\text{B}}$ είναι διαφορετική της συχνότητας άφιξης των κλάσεων C_{Γ}, C_{Δ}. Συγκεκριμένα στο πείραμα 1 ο ρυθμός άφιξης των κλάσεων δοσοληψιών C_{A}, C_{B} μεταβάλλεται στη διάρκεια του χρόνου προσομοίωσης ως εξής:

Ρυθμός Άφιξης : $\frac{10}{15}[0, 500]$
$\frac{10}{10}[500, 1000]$
$\frac{10}{9}[1000, 2000]$
$\frac{10}{8}[2000, 2500]$
$\frac{10}{7}[2500, 3600]$

ενώ ο ρυθμός άφιξης των κλάσεων C_{Γ}, C_{Δ} είναι μεγαλύτερος σε κάθε διάστημα του χρόνου προσομοίωσης.

Ρυθμός Άφιξης : $1.0[0, 500]$
$\frac{10}{9}[500, 1000]$
$\frac{10}{8}[1000, 2000]$
$\frac{10}{7}[2000, 2500]$
$\frac{10}{6}[2500, 3600]$

Τα αποτελέσματα του πειράματος προσομοίωσης 1 δείχνουν ότι οι κλάσεις δοσοληψιών C_{A}, C_{B} που παρουσιάζουν μικρότερη συχνότητα εμφάνισης σε όλους τους κόμβους του συστήματος επιτυγχάνουν μικρότερους μέσους χρόνους απόκρισης συγκριτικά με τις κλάσεις C_{Γ}, C_{Δ}. Η κλάση C_{A} παρουσιάζει μικρότερο μέσο χρόνο απόκρισης σε σχέση

με την κλάση C_B (με την οποία παρουσιάζει την ίδια συχνότητα εμφάνισης) αφού η σχέση Α που προσπελαύνει έχει αντιγραφεί σε περισσότερους κόμβους. Για τον ίδιο λόγο η κλάση C_Δ παρουσιάζει μικρότερο μέσο χρόνο απόκρισης από την κλάση C_Γ. Παρατηρούμε επίσης ότι οι σχέσεις Γ, Δ που παρουσιάζουν μεγαλύτερη ζήτηση έχουν συνολικά αντιγραφεί σε περισσότερους κόμβους. Στο πείραμα προσομοίωσης 2 οι κλάσεις δοσοληψιών C_A, C_B είναι αυτές που παρουσιάζουν μεγαλύτερη συχνότητα άφιξης σε όλους τους κόμβους του συστήματος. Τα αποτελέσματα του πειράματος 2 επιβεβαιώνουν τα συμπεράσματα που έχουν εξαχθεί από το πείραμα 1. Οι κλάσεις δοσοληψιών C_Γ, C_Δ που παρουσιάζουν μικρότερο ρυθμό άφιξης επιτυγχάνουν μικρότερους μέσους χρόνους απόκρισης. Η κλάση C_Δ παρουσιάζει μικρότερο μέσο χρόνο απόκρισης των δοσοληψιών της σε σχέση με την κλάση C_Γ. Αυτό οφείλεται στο ότι η σχέση Δ που προσπελαύνεται από τις δοσοληψίες της κλάσης C_Δ έχει αντιγραφεί σε περισσότερους κόμβους. Ανάλογα η κλάση C_A παρουσιάζει μικρότερο μέσο χρόνο απόκρισης των δοσοληψιών της σε σχέση με την κλάση C_B (η σχέση Α έχει αντιγραφεί σε περισσότερους κόμβους). Στο πείραμα 2 οι σχέσεις Α, Β που παρουσιάζουν μεγαλύτερη ζήτηση έχουν αντιγραφεί συνολικά σε περισσότερους κόμβους.

Αν δεν επιτρεπόταν η αντιγραφή των σχέσεων της βάσης δεδομένων στους κόμβους του συστήματος οι κλάσεις C_A, C_B, C_Γ, C_Δ θα παρουσίαζαν χειρότερους μέσους χρόνους απόκρισης των δοσοληψιών τους.

	$RespC_A$	$RespC_\mathrm{B}$	$RespC_\Gamma$	$RespC_\Delta$
DRR	0.38	0.38	0.31	0.41
ECO/COMM/SQRD	0.46	0.46	0.44	0.45

Πίνακας 4

Στον Πίνακα 4 σημειώνονται τα αποτελέσματα πειραμάτων προσομοίωσης, στα οποία οι κλάσεις δοσοληψιών παρουσιάζουν κάθε χρονική στιγμή τον ίδιο ρυθμό άφιξης σε όλους τους κόμβους του συστήματος και οι σχέσεις Α, Β, Γ, Δ έχουν ανατεθεί στους κόμβους P_1, P_2, P_3, P_4 αντίστοιχα. Στην πρώτη γραμμή παρουσιάζονται τα αποτελέσματα προσομοίωσης όταν ο αλγόριθμος δρομολόγησης είναι ο DRR ενώ στη δεύτερη όταν απλώς κάθε δοσοληψία δρομολογείται στον κόμβο που διαθέτει τη σχέση που επιθυμεί να προσπελάσει. Παρατηρούμε ότι ο αλγόριθμος DRR είναι αποδοτικότερος αφού επιτυγχάνει μικρότερους μέσους χρόνους απόκρισης για όλες τις κλάσεις δοσοληψιών.

	$RespC_A$	$RespC_{\text{B}}$	$RespC_{\Gamma}$	$RespC_{\Delta}$
ECO	0.43	0.43	0.45	0.46
COMM	0.62	0.51	0.60	0.54
SQRD	0.43	0.43	0.45	0.46

Πίνακας 5

Στον Πίνακα 5 παρουσιάζονται τα αποτελέσματα πειραμάτων προσομοίωσης στα οποία έχει χρησιμοποιηθεί το σύστημα επεξεργασίας δοσοληψιών του Πίνακα 4 αλλά η ανάθεση των δεδομένων είναι διαφορετική. Η σχέση Α έχει αντιγραφεί στους κόμβους P_1, P_2 ενώ η σχέση Β έχει αντιγραφεί στους κόμβους P_2, P_3. Παρατηρούμε ότι οι αλγόριθμοι ECO,COMM,SQRD δίδουν μεγαλύτερους χρόνους απόκρισης όλων των κλάσεων δοσοληψιών συγκρινόμενοι με τον αλγόριθμο DRR (του Πίνακα 4).

	$RespC_A$	$RespC_{\text{B}}$	$RespC_{\Gamma}$	$RespC_{\Delta}$
ECO	0.42	0.42	0.42	0.43
COMM	0.58	0.49	0.54	0.52
SQRD	0.43	0.42	0.43	0.44

Πίνακας 6

Στον Πίνακα 6 παρουσιάζονται τα αποτελέσματα προσομοιώσεων στις οποίες έχει χρησιμοποιηθεί το ίδιο μοντέλο συστήματος αλλά έχει αυξηθεί το ποσοστό αντιγραφής των σχέσεων στους κόμβους του συστήματος. Η σχέση Α έχει αντιγραφεί στους κόμβους P_1, P_2, η σχέση Β στους κόμβους P_2, P_3, η σχέση Γ στους κόμβους P_1, P_3, ενώ η σχέση Δ έχει ανατεθεί μόνο στον κόμβο P_4. Παρατηρούμε ότι οι οικονομικοί αλγόριθμοι ECO, COMM, SQRD εξακολουθούν να δίνουν μεγαλύτερους μέσους χρόνους απόκρισης όλων των κλάσεων δοσοληψιών σε σχέση με τον αλγόριθμο DRR (Πίνακας 4) παρά την αύξηση αντιγραφής των σχέσεων της βάσης.

Αν το ποσοστό δεδομένων που μπορείς να βρεις σε ένα κόμβο συνεχίζει να αυξάνει οι οικονομικοί αλγόριθμοι ECO,COMM,SQRD θα δώσουν καλύτερα αποτελέσματα από τον αλγόριθμο DRR του Πίνακα 4. Στον Πίνακα 7 σημειώνονται τα αποτελέσματα πειραμάτων προσομοίωσης στα οποία οι σχέσεις Β, Δ έχουν αντιγραφεί σε όλους τους κόμβους του συστήματος.

	$RespC_A$	$RespC_\text{B}$	$RespC_\Gamma$	$RespC_\Delta$
ECO	0.28	0.27	0.27	0.29
COMM	0.45	0.34	0.40	0.37
SQRD	0.28	0.27	0.28	0.30

Πίνακας 7

Στις προσομοιώσεις αυτές παρατηρείται ότι η δρομολόγηση δοσοληψιών με βάση τους οικονομικούς αλγορίθμους ECO, COMM, SQRD κατέληξε σε μικρότερους μέσους χρόνους απόκρισης όλων των κλάσεων δοσοληψιών σε σχέση με τον αλγόριθμο DRR. Αυτό οφείλεται στο ότι η ζήτηση και η προσφορά του συστήματος που χρησιμοποιήθηκε στον Πίνακα 4 δεν επέτρεψε την αντιγραφή των σχέσεων σε όλους τους κόμβους (η σχέση Α είχε αντιγραφεί στους κόμβους P_1, P_3, P_4, η σχέση Β στους κόμβους P_1, P_2, P_3, ενώ η σχέση Δ είχε αντιγραφεί στους κόμβους P_1, P_2, P_4). Ακόμη και αν είχε επιτραπεί η αντιγραφή των σχέσεων σε όλους τους κόμβους αυτό δεν θα επιτυγχανόταν από το πρώτο διάστημα του χρόνου προσομοίωσης με αποτέλεσμα οι αλγόριθμοι ECO, COMM, SQRD να διατηρούσαν το προβάδισμα.

Ο αλγόριθμος DRR συγκρίθηκε επίσης με τους αλγορίθμους ικανοποίησης στόχων επίδοσης **WFW** (Wait For Work), **WFWC** (Wait For Work Class), **SGOR** (Simple Goal Oriented Algorithm) που έχουν υλοποιηθεί από τους Ferguson *et. al.* [9].

- **Αλγόριθμος WFW**

  Ο αλγόριθμος επιδιώκει να ελαχιστοποιήσει το μέσο χρόνο απόκρισης κάθε δρομολογούμενης δοσοληψίας. Εστω Τ δοσοληψία της κλάσης C_i. Ο αλγόριθμος υπολογίζει τον αναμενόμενο χρόνο απόκρισης αν η δοσοληψία δρομολογηθεί σε κάθε σύστημα S_j. Ο χρόνος αυτός είναι συνάρτηση της υπολογιστικής καθυστέρησης (CPU delay), της καθυστέρησης του δικτύου (network delay), της καθυστέρησης στις συσκευές εισόδου/εξόδου (I/O delay) και της καθυστέρησης για ενημέρωση του ημερολογίου (logging delay)

- **Αλγόριθμος WFWC**

  Ο αλγόριθμος αυτός αποτελεί μία επέκταση του αλγόριθμου WFW που λαμβάνει επιπλέον υπόψη στην απόφαση δρομολόγησης την προτεραιότητα της δρομολογούμενης δοσοληψίας και την προτεραιότητα των δοσοληψιών που υπάρχουν ήδη στο σύστημα.

- **Αλγόριθμος SGOR**

  Ο αλγόριθμος προβλέπει την επίδραση κάθε πιθανής δρομολόγησης της δοσοληψίας στην ικανοποίηση στόχων όλων των κλάσεων δοσοληψιών και δρομολογεί τη δοσοληψία στο σύστημα που εγγυάται καλύτερη ικανοποίηση των στόχων όλων των κλάσεων.

Ολα τα πειράματα προσομοίωσης που διεξήχθηκαν για να συγκριθεί η συμπεριφορά αυτών των αλγορίθμων με τον οικονομικό αλγόριθμο DRR επέδειξαν ως αποτελεσματικότερο τον αλγόριθμο DRR. Παράδειγμα αυτών των πειραμάτων δίδεται στον Πίνακα 8

	$RespC_A$	$RespC_\mathrm{B}$	$RespC_\Gamma$	$RespC_\Delta$
DRR	0.86	0.99	0.77	0.63
WFN	0.98	1.02	0.94	0.92
WFWC	0.98	1.02	0.94	0.92
SGOR	0.90	0.85	0.98	0.84

Πίνακας 8

Στο μοντέλο που χρησιμοποιήθηκε (Σχήμα 4.2) οι σχέσεις Α, Β, Γ, Δ έχουν ανατεθεί στους κόμβους P_1, P_2, P_3, P_4 αντίστοιχα.

4.5 Επισκόπηση Σχετικής Βιβλιογραφίας

Το θέμα της ανάθεσης αρχείων δεδομένων (files) σε ένα πολύπλοκο υπολογιστικό σύστημα έχει μελετηθεί από τον Kurose *et. al.* [11]. Η εργασία τους παρέχει μη κεντρικοποιημένους αλγόριθμους για κατανομή πόρων (files, file fragments) μεταξύ συνεργαζομένων πρακτόρων (υπολογιστικό σύστημα).Το κόστος επικοινωνίας και η αναμενόμενη καθυστέρηση εξυπηρέτησης περιλαμβάνονται στο κριτήριο βελτιστοποίησης (utility function) το οποίο είναι κοινό για όλους τους χρήστες (κόμβους) του κατανεμημένου συστήματος.

Το πρόβλημα της μετανάστευσης και αντιγραφής δεδομένων έχει μελετηθεί σε οικονομικό περιβάλλον και από τους Ferguson *et. al.* [6, 4, 7]. Οι δοσοληψίες που εισέρχονται στο σύστημα για εξυπηρέτηση, χρεώνονται από τους επεξεργαστές (προμηθευτές) του συστήματος για τις πράξεις ανάγνωσης, εγγραφής στα δεδομένα της βάσης. Οι επεξεργαστές ενοικιάζουν πόρους (δεδομένα) σε άλλους επεξεργαστές

του συστήματος, ώστε να αυξήσουν τα προσωπικά τους κέρδη από τα έσοδα των ενοικιάσεων. Το πρόβλημα της διαχείρισης δεδομένων ορίζεται και στις εργασίες αυτές ως ελαχιστοποίηση του μέσου χρόνου απόκρισης των δοσοληψιών.

Πρόσφατα ο Stonebraker*et. al.* μελέτησε το πρόβλημα της εκτέλεσης επερωτήσεων και της μετανάστευσης δεδομένων σε μεγάλα κατανεμημένα συστήματα που αποτελούνται από χιλιάδες κόμβους. Εχουν επιλέξει ένα οικονομικό παράδειγμα παρόμοιο με αυτό που περιγράφηκε στις εργασίες [5, 6, 8] για την σχεδίαση μιας κατανεμημένης βάσης δεδομένων, που ονομάζουν Mariposa [23]. Εχουν σχεδιάσει επαρκής μηχανισμούς πλειστηριασμού για την προσπέλαση των δεδομένων από τους πελάτες (χρήστες) του συστήματος και την κατάλληλη αντιγραφή τους στους κόμβους (προμηθευτές) του συστήματος.

Κεφάλαιο 5

Συμπεράσματα και Ερευνητικές Κατευθύνσεις

5.1 Σύνοψη Αποτελεσμάτων

Σε αυτή την εργασία μελετήθηκε η εφαρμογή οικονομικών αλγορίθμων για την κατανομή φόρτου εργασιών και τη διαχείριση δεδομένων σε κατανεμημένα συστήματα. Η κατανομή φόρτου εργασιών επιτεύχθηκε με τον σχεδιασμό μίας οικονομίας όπου οι προμηθευτές (κόμβοι του συστήματος) επιδιώκουν την μεγιστοποίηση των κερδών τους διαθέτοντας τα προϊόντα τους (υπολογιστικό χρόνο, εύρος επικοινωνίας) σε καταναλωτές (δοσοληψίες). Στόχος των καταναλωτών είναι η εξυπηρέτησή τους με το μικρότερο εφικτό κόστος. Η διαχείριση δεδομένων επιτεύχθηκε με τον σχεδιασμό μίας οικονομίας στην οποία οι καταναλώτες (δοσοληψίες) πληρώνουν τους προμηθευτές (κόμβοι του συστήματος) για την εκτέλεση πράξεων ανάγνωσης στα τοπικά τους δεδομένα, για την αγορά υπολογιστικού χρόνου και εύρους επικοινωνίας. Οι προμηθευτές ενοικιάζουν δεδομένα που δε διαθέτουν τοπικά όταν και τα δύο μέλη της πράξης ενοικίασης κρίνουν οτι μακροπρόθεσμα θα επωφεληθούν οικονομικά (ο ενοικιαστής θα αποκομίσει περισσότερα κέρδη από την εξυπηρέτηση δοσοληψιών που ζητούν το καινούριο αντίγραφο, ενώ ο ιδιοκτήτης θα εισπράτει το ενοίκιο και θα αντισταθμίζει τη πιθανή ζημιά που μπορεί να προκαλέσει ένα επιπλέον αντίγραφο). Στόχος των προμηθευτών είναι η μεγιστοποίηση των προσωπικών τους κερδών, ενώ των καταναλωτών να εξυπηρετηθούν με το μικρότερο εφικτό κόστος.

Γνωρίζουμε ότι το αυξανόμενο μέγεθος και η δυναμική συμπεριφορά των σύγχρονων

υπολογιστικών συστημάτων καθώς και η ετερογένεια των χρηστών που χρησιμοποιούν τα συστήματα αυτά καθιστά ανεπαρκείς τις παραδοσιακές μεθόδους για κατανομή πόρων. Στα συστήματα αυτά πολλές φορές είναι αδύνατος ο καθορισμός ενός καθολικού στόχου επίδοσης (perfomance goal) που αντιπροσωπεύει τις απαιτήσεις και τους στόχους των ετερογενών χρηστών, αλλά ακόμη και αν ένας τέτοιος στόχος οριστεί είναι δύσκολο να επιτευχθεί σε ένα τόσο δυναμικό περιβάλλον. Η μοντελοποίηση του συστήματος ως οικονομία επιλύει το πρόβλημα. Κάθε πράκτορας της οικονομίας (καταναλωτής, προμηθευτής) έχει ένα προσωπικό στόχο και κανόνες για την επίτευξη του στόχου αυτού. Δεν υπάρχει ανάγκη ορισμού ένος καθολικού στόχου επίδοσης με αποτέλεσμα τη μείωση της πολυπλοκότητας της διαδικασίας κατανομής των πόρων του συστήματος. Η επίτευξη των επιμέρους στόχων επίδοσης των πρακτόρων οδηγεί σε μία καθολικά ικανοποιητική ανάθεση των πόρων του συστήματος.

Στην εργασία αυτή αποδείχθηκε επίσης οτι οι οικονομικοί αλγόριθμοι μπορούν να επιλύσουν το πρόβλημα της κατανομής πόρων το ίδιο καλά με μη οικονομικούς αλγόριθμους. Η σύγκριση του οικονομικού αλγόριθμου DRR με τους αλγόριθμους ικανοποίησης στόχων επίδοσης WFW, WFWC και SGOR δείχνει οτι όλες οι κλάσεις δοσοληψιών των υπό μελέτη συστημάτων επιτυγχάνουν με τον αλγόριθμο DRR μέσο χρόνο απόκρισης μικρότερο από τους αντίστοιχους χρόνους που προκύπτουν με τους άλλους αλγορίθμους.

Εχει αποδειχθεί επίσης οτι οι οικονομικοί αλγόριθμοι χειρίζονται ικανοποιητικά την ύπαρξη των διαφορετικών πόρων του συστήματος (υπολογιστικό χρόνο, εύρος επικοινωνίας, δεδομένα) και επιτρέπουν την ικανοποίηση χρηστών με ετερογενείς προτιμήσεις (διαφορετικές συναρτήσεις χρησιμότητας). Οι τιμές των πόρων καθρεπτίζουν τη ζήτηση και την προσφορά τους στο σύστημα και διαμορφώνονται με βάση τις υπολογιστικές δυνατότητες και το φόρτο του συστήματος καθώς και τις απαιτήσεις των χρηστών. Η τιμή ενός πόρου επηρεάζει τη συμπεριφορά του συστήματος και την δρομολόγηση των δοσοληψιών. Για παράδειγμα αν το κόστος επικοινωνίας είναι υψηλό και έχει μεγάλο βάρος στην απόφαση δρομολόγησης των εργασιών (αλγόριθμος COMM) θα περιοριστεί η μετανάστευση δοσοληψιών στους κόμβους του συστήματος. Στην εργασία παρατηρήθηκε οτι καλύτερη απόδοση (ελαχιστοποίηση του μέσου χρόνου απόκρισης των κλάσεων δοσοληψιών) επιτυγχάνεται όταν το κόστος εξυπηρέτησης (κόστος αναμονής + κόστος εκτέλεσης) έχει μεγαλύτερο βάρος στην απόφαση δρομολόγησης από το κόστος

επικοινωνίας (αλγόριθμοι ECO, SQRD, DRR).

Η μελέτη της συμπεριφοράς του οικονομικού αλγόριθμου διαχείρισης δεδομένων DRR έδειξε οτι το πλήθος των αντιγράφων ενός δεδομένου εξαρτάται από τον φόρτο του συστήματος και τη ζήτηση του δεδομένου (πλήθος των πράξεων ανάγνωσης). Τα αντίγραφα αυτά δημιουργούνται σε κόμβους στους οποίους παρατηρείται μεγάλη ζήτηση για το δεδομένο αυτό. Η οικονομική αυτή πολιτική είναι προτιμότερη από την διατήρηση μοναδικού αντιγράφου ενός δεδομένου σε ένα μόνο κόμβο (no replication) αφού επιτυγχάνει μικρότερους χρόνους απόκρισης των δοσοληψιών. Είναι επίσης προτιμότερη από την διατήρηση αντιγράφου κάθε δεδομένου σε όλους τους κόμβους του συστήματος (total repliation) αφού επιτρέπει τη δημιουργία αντιγράφων μόνο αν υπάρχει ανάλογη ζήτηση περιορίζοντας έτσι σημαντικά το κόστος αποθήκευσης (εξοικονόμηση χώρου στους ενταμιευτές του συστήματος).

Στον σχεδιασμό της οικονομίας τόσο για την κατανομή του φόρτου εργασιών όσο και για τη διαχείριση δεδομένων θεωρήσαμε οτι κάθε δοσοληψία που εισέρχεται στο σύστημα έχει ‘‘a priory’’ γνώση των υπολογιστικών της απαιτήσεων των πόρων δηλαδή που χρειάζεται για να εκτελεστεί. Αν και στα οικονομικά συστήματα είναι τυπικό να θεωρήσουμε οτι κάθε πράκτορας γνωρίζει τις απαιτήσεις του σε αγαθά η θεώρηση αυτή δεν είναι ρεαλιστική σε υπολογιστικά συστήματα και αποτελεί μειονέκτημα της παρούσας εργασίας.

Ενα ακόμη μειονέκτημα είναι οτι οι πόροι ενός υπολογιστικού συστήματος είναι αδιαίρετοι γεγονός που μπορεί να οδηγήσει σε μη συνεχής συναρτήσεις χρησιμότητας. Είναι λοιπόν πολύ πιθανό οι οικονομικοί αλγόριθμοι να βρουν υποβέλτιστη λύση στο πρόβλημα της κατανομής των πόρων του συστήματος να υπάρχει δηλαδή ανάθεση πόρων που οδηγεί σε μικρότερους μέσους χρόνους απόκρισης των κλάσεων δοσοληψιών.

5.2 Επεκτάσεις και Ερευνητικές Κατευθύνσεις

Στην οικονομία διαχείρισης δεδομένων οι δοσοληψίες που εισέρχονται στο σύστημα υποβάλουν μόνο πράξης ανάγνωσης στους κόμβους του συστήματος. Αν επιτρεπόταν να υποβληθούν και πράξεις εγγραφής ο αριθμός των αντιγράφων κάθε δεδομένου θα εξαρτιόταν από το λόγο του αριθμού των πράξεων ανάγνωσης και του αριθμού των πράξεων εγγραφής για το δεδομένο αυτό. Αν ο αριθμός των πράξεων εγγραφής για ένα

δεδομένο ήταν μεγάλος τα αντίγραφα του δεδομένου αυτού θα ήταν περιορισμένα για να μειωθεί το κόστος ενημέρωσης. Αν αντίθετα ο αριθμός των πράξεων ανάγνωσης για ένα δεδομένο ήταν σχετικά μεγαλύτερος θα αυξανόταν η αντιγραφή του δεδομένου αυτού ώστε οι δοσοληψίες να είναι δυνατό να εκτελούνται τοπικά σε κάθε κόμβο ή σε γειτονικούς κόμβους. Στις μελλοντικές προεκτάσεις αυτής της εργασίας περιλαμβάνεται η υλοποίηση ενός πρωτοκόλλου ενημέρωσης των αντιγράφων και ελέγχου συνέπειας της βάσης δεδομένων ώστε να είναι εφικτή και η υποβολή αιτήσεων εγγραφής από τις δοσοληψίες του συστήματος.

Στις ερευνητικές κατευθύνσεις αυτής της εργασίας περιλαμβάνεται επίσης και η μελέτη πιο πολύπλοκων οικονομικών αλγορίθμων που θα λαμβάνουν υπόψη περισσότερες παραμέτρους απόδοσης των υπολογιστικών συστημάτων (π.χ κόστος αποθήκευσης) και θα χρησιμοποιούν διαφορετικούς μηχανισμούς για την κοστολόγηση των αγαθών και των υπηρεσιών (π.χ πλειστηριασμός).

Τα οικονομικά μοντέλα θα μπορούσαν επίσης να χρησιμοποιηθούν για την παροχή υπηρεσιών στο **Internet** το μέγεθος του οποίου συνεχώς αυξάνει. Αυτή τη στιγμή το Internet συνδέει εκατομύρια χρήστες και εκαντοντάδες κόμβους. Σε πολλούς από αυτούς τους κόμβους υπάρχουν υπηρεσίες όπως το **World Wide Web (WWW)** που παρέχουν πρόσβαση σε διάφορες πηγές πληροφοριών. Πολλές επιπλέον υπηρεσίες όπως για παράδειγμα πολύμορφες εφαρμογές (multimedia applications) αναμένεται να υποστηρικτούν στο μέλλον από το Internet. Οι χρήστες του δικτύου θα πρέπει να μοιράζονται περιορισμένο εύρος επικοινωνίας(network bandwidth) και υπολογιστική δύναμη για να έχουν πρόσβαση σε αυτές τις υπηρεσίες. Σε τόσο μεγάλα δίκτυα είναι απαραίτητη η εύρεση ενός μη κεντρικοποιημένου μηχανισμού για τον έλεγχο πρόσβασης στις παρεχόμενες υπηρεσίες. Οι οικονομικοί αλγόριθμοι παρέχουν ένα μη κεντρικοποιημένο μηχανισμό κατανομής πόρων και έτσι μπορούν να μειώσουν σημαντικά την πολυπλοκότητα της διαδικασίας παροχής υπηρεσιών στο Internet. Στις μελλοντικές προεκτάσεις αυτής της εργασίας περιλαμβάνεται η εφαρμογή οικονομικών αλγοριθμών για κατανομή πόρων στο Internet.

Βιβλιογραφία

[1] K. Arrow *et. al.*, *Handbook of Mathematical Economics*, vol I and vol II, North Holland Publishing, Amsterdam, 1981.

[2] G. Debreu, *The Theory of Value*, NewYork, Wiley, 1959.

[3] Y. Yemini, ''Selfish Optimization in Computer Networks,'' *Proceedings of the 20th IEEE Conf. on Decision and Control*, pages 281-285, San Diego, Decemeber, 1981.

[4] D. Ferguson, C. Nikolaou, J. Sairamesh, Y. Yemini, ''Economic Models for Resource Allocation in Computer Systems,'' *in Market-based Control: A Paradigm for Distributed Resource Allocation*, ed. Scott H. Clearwater, World Scientific Press, Singapore, 1995.

[5] D.F. Ferguson, C. Nikolaou and Y. Yemini, ''An Economy for Managing Replicated Data in Autonomous Decentralized Systems,'' *Proc. Int. Symp. on Autonomous Decentralized Sys. (ISADS'93)*, Kawasaki, Japan, 1993.

[6] D.F. Ferguson, *The application of Microeconomics to the design of resourse allocation and control algorithms in Distributed Systems*, PhD thesis, Columbia University, New York, 1989.

[7] D.F. Ferguson, Y. Yemini and C. NiKolaou, ''Microeconomic Algorithms for Load Balancing in Distributed Computer Systems,'' *Proc. of International Conference on Distributed Systems (ICDCS'88).*

[8] D.F. Ferguson, Y. Yemini, C. Nikolaou, ''An Economy for Flow Control in Computer Networks,'' *Proc. of the INFOCOM'90.*

[9] D.F. Ferguson, L. Georgiadis, C. Nikolaou, K.Davies ''Satisfying response time goals in transaction processing systems,'' *Technical Report RC 18139, IBM T.J. Watson Research Center*, 1992.

[10] J.F. Kurose, M. Schwartz and Y. Yemini, ''A microeconomic approach to optimization of channel access policies in multiaccess networks,'' *Proc. of 5th Int. Conf. Distrib. Comput. Syst.*, Denver, May 1985.

[11] J.F. Kurose and R. Simha, ''A Microeconomic approach to optimal resource allocation in distributed computer systems,'' *IEEE Trans. on Comp.*, May 1989.

[12] L. Kleinrock, *Queueing Systems: Volume 1: Theory*, Wiley Interscience, NewYork, 1975.

[13] H. R. Varian, *Microeconomic Analysis*, 3rd ed., Norton Press, 1992.

[14] W. Hildenbrand and A.P. Kirman, *Introduction to Equilibrium Analysis*, North Holland Publishing Co., Amsterdam, 1976.

[15] M. Miller and E. Drexler, ''Markets and Computation: Agoric Open Systems,'' *The Ecology of Computation*, North Holland, Amsterdam, 1988.

[16] T.W. Malone, R.E. Fikes, K.R. Grant and M.T. Howard, ''Enterprice: A Market-Like Task Scheduler for Distributed Computing Environments,'' *The Ecology of Computation*, North Holland, Amsterdam, 1988.

[17] R. G Lipsey and P.O Steiner, *Microeconomics*, Harper and Row, New York, 1979.

[18] J. Sairamesh, D. Ferguson, and Y. Yemini, ''An Approach to Pricing, Optimal Allocation and Quality of Service Provisioning in High Speed Packet Networks,'' *Proceedings of the INFOCOM'95*.

[19] A. Bovopoulous and A. A. Lazar, ''Decentralized Algorithms for Optimal Flow Control,'' *Proc. of the 25th Annual Allerton Conference on Commuinication, Control and Computing*, September 1987.

[20] M.T. Hsiao and A.A. Lazar, *Optimal Decentralized Flow Control in Computer Communication Networks*, Phd thesis, CTR-Tech report, Columbia University, NY, 1986.

[21] C. Dougleris and R. Mazumdar, ''On Pareto Optimal Flow Control in a Multiclass Environment,'' *CTR Technical Report, Columbia University*, 1987.

[22] M. Marazakis, ''Simulation of TP Systems and a Study of Methods for Perfomance Goal Satisfaction,'' *Technical Report 140, Dep. of Computer Science, University of Crete*, 1996.

[23] M. Stonebraker *et. al.*, ''An Economic Paradigm For Query Processing and Data Migration in Mariposa,'' *Proc. of Parallel and Distributed Information Systems*, pages 58-67, 1994.

[24] R. Cocchi *et. al.*, ''A Study of priority pricing in multiclass networks,'' *Proc. of the SIGCOMM'91.*

[25] M. Livny, *The Study of Load Balancing Algorithms for Decentralized Distributed Processing Systems*, PhD thesis, Weizmann Institute of Science, Rehovot, Israel, 1984.

[26] C. A. Waldspurger, T. Hogg, B. Huberman, J. Kephart, and S. Stornetta, ''Spawn: A distributed computational ecology,'' *IEEE Trans. on Software Engg.*, 18(2), February, 1992.

[27] C. A. Waldspurger and William E. Weihl, ''Lottery Scheduling: Flexible Proportional-Share Resource Mangement,'' *Proceedings of the First Symposium on Operating Systems Design and Implementation*, November 1994.

[28] Carl A. Waldspurger and William E. Weihl, ‘‘Stride Scheduling: Deterministic Proportional-Share Resource Mangement,’’ *Technical MemMIT/LCS/TM-528*, June 1995.

[29] A. T. Bouloutas and D. N. Serpanos, ‘‘A comparision of Multimedia Architectures,’’ *IBM Research Report--RC19162, 1993*

[30] L. Golubchik, J.C.S. Lui and R. Mutz, ‘‘Reducing I/O Demand in Video-0n-Demand Storage Servers,’’ *SIGMETRICS*, 1995.