ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ

ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ

Ανάλυση απόδοσης και αξιοπιστίας για πολιτικές πλεονασματικών αντικειμένων σε συστήματα ανοχής λαθών

Διπλωματική Εργασία

των:

Ιακωβίδου Νάντια και **Θεόδωρου Σολδάτου**

Επιβλέπων καθηγητής: κ. Παναγιώτης Κατσαρός

Θεσσαλονίκη 2005

ΠΕΡΙΕΧΟΜΕΝΑ

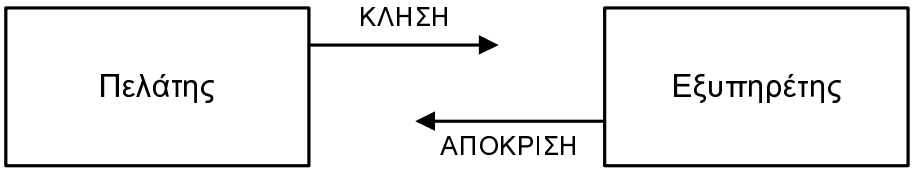
ΚΕΦΑΛΑΙΟ 1ο

Εισαγωγή

1.1 Βασικές Έννοιες

Η εργασία αυτή παρουσιάζει μια προσέγγιση αξιολόγησης βασισμένη σε προσομοίωση που στόχο έχει την αντιστάθμιση των παραγόντων που επηρεάζουν την απόδοση και την αποτελεσματικότητα της ανοχής λαθών σε συστήματα πλεονασματικής επεξεργασίας (replication schemes). Στα συστήματα αυτά γίνεται χρήση πιθανώς διαφορετικών πολιτικών πλεονασματικής επεξεργασίας στα αντικείμενα που τα απαρτίζουν (*σύνθετα σχήματα πλεονασματικής επεξεργασίας*). Η προαναφερόμενη προσέγγιση χρησιμοποιείται στην ανάλυση της *απόδοσης* μιας περίπτωσης υποθετικού *συστήματος* που μας έδωσε την ευκαιρία να μελετήσουμε έναν αριθμό από σύνθετες *πολιτικές ανοχής λαθών*.

Το *κατανεμημένο σύστημα* που μελετήσαμε, βασίζεται στο μοντέλο πελάτη-εξυπηρέτη. Η βασική ιδέα πίσω από το μοντέλο αυτό είναι ότι το σύστημα αποτελείται από ένα σύνολο αντικειμένων, τα οποία αλληλεπιδρούν μεταξύ τους με σκοπό να εξυπηρετήσουν ένα σύνολο από κλήσεις (requests). Τα αντικείμενα που εξυπηρετούν τις κλήσεις ονομάζονται εξυπηρέτες και τα αντικείμενα που στέλνουν κλήσεις προς εξυπηρέτηση παίζουν το ρόλο των πελατών (client objects).

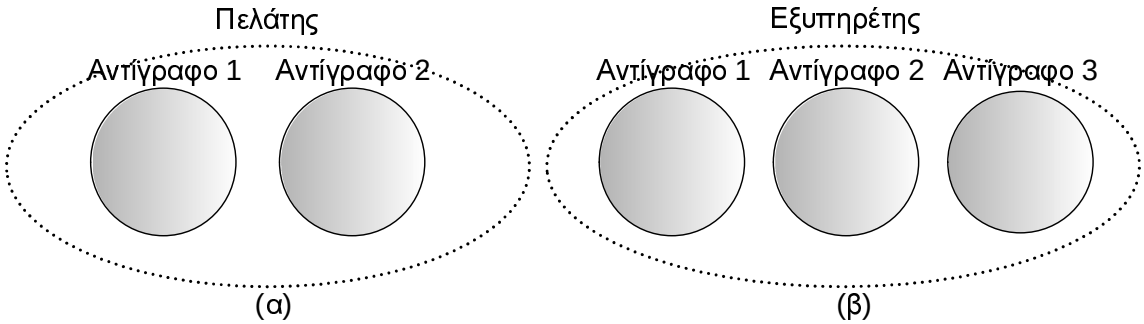


Σχήμα 1. Παράδειγμα επικοινωνίας πελάτη – εξυπηρέτη.

Το ζήτημα της *απόδοσης* για ένα κατανεμημένο σύστημα βρίσκεται κάθε χρονική στιγμή στο προσκήνιο. Για τη μέτρηση της απόδοσης χρησιμοποιούνται διάφορα μέτρα, ένα από τα οποία είναι ο ρυθμός απόδοσης (throughput) (αριθμός κλήσεων που εξυπηρετούνται ανά μονάδα χρόνου).

Ένα σημαντικό πλεονέκτημα των κατανεμημένων συστημάτων είναι ότι μπορούν να παρέχουν υψηλή *αξιοπιστία* (*reliability*). Αυτό σημαίνει ότι επειδή υπάρχει κατανομή του φόρτου εργασίας σε πολλούς εξυπηρέτες, η διακοπή λειτουργίας ενός εξυπηρέτη δε θα προκαλέσει και την κατάρρευση του συστήματος ως ολότητα. Αντιθέτως το σύστημα θα συνεχίσει να λειτουργεί. Η υψηλή αξιοπιστία αποτελεί καθοριστικό παράγοντα ιδιαίτερα σε κρίσιμες εφαρμογές όπως για παράδειγμα στον έλεγχο πυρηνικών αντιδραστήρων, ιατρικών εφαρμογών, κτλ.

Μια όψη της αξιοπιστίας είναι η διαθεσιμότητα (availability), η οποία αναφέρεται στο ποσοστό του χρόνου κατά το οποίο το σύστημα είναι διαθέσιμο. Ένα εργαλείο βελτίωσης της διαθεσιμότητας είναι ο *πλεονασμός* (*redundancy*). Θα πρέπει δηλαδή ο εξυπηρέτης να αποτελείται από πολλά αντίγραφα του εαυτού του (*πλεονασματικά αντικείμενα*) (*replicas*) έτσι ώστε αν διακοπεί η λειτουργία κάποιου εξ αυτών, τα υπόλοιπα να είναι σε θέση να αναπληρώσουν το κενό που εμφανίζεται. Στο σύνολο των αντιγράφων που συνιστούν έναν εξυπηρέτη θα αναφερόμαστε με τον όρο *ομάδα αντικειμένων* (*object group*). Εκτός από έναν εξυπηρέτη βέβαια, ένα σύνολο αντιγράφων μπορεί επίσης να συνιστά έναν πελάτη (στο Σχήμα 2 φαίνονται κάποια παραδείγματα ομάδας αντικειμένων). Ένα άλλο ζήτημα που σχετίζεται με την αξιοπιστία είναι η *ανοχή σε λάθη* (*fault tolerance*). Εάν δηλαδή κάποιο από τα αντίγραφα ενός αντικειμένου καταρρεύσει να μπορεί να γίνει *επαναφορά* (*recovery*) αυτού έτσι ώστε να μπορεί να συνεχίσει τη λειτουργία του.



Σχήμα 2. (α) Ομάδα αντικειμένων ενός πελάτη. (β) Ομάδα αντικειμένων ενός εξυπηρέτη.

1.2 Είδη Σφαλμάτων (Faults)

Σε αυτήν την παράγραφο θα αναφέρουμε κάποια από τα πιθανά σφάλματα που μπορούν να συμβούν σε ένα κατανεμημένο σύστημα που βασίζεται στο μοντέλο πελάτη - εξυπηρέτη.

1. Είναι πιθανό ο πελάτης να μην μπορεί να εντοπίσει την ύπαρξη εξυπηρέτη. Στην περίπτωση αυτή, ο εξυπηρέτης είναι πολύ πιθανό να είναι εκτός λειτουργίας.

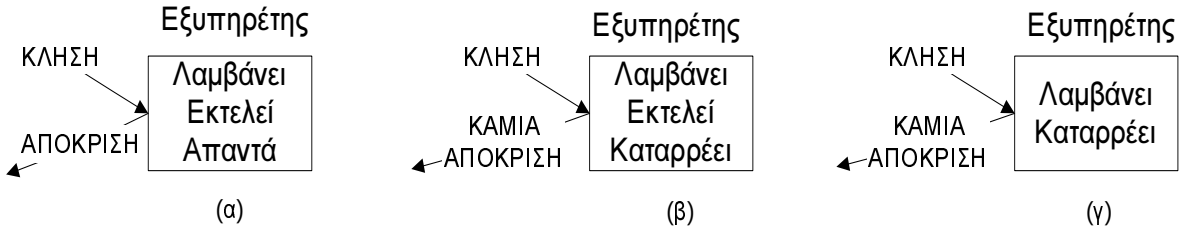
2. Μια άλλη περίπτωση είναι να χαθεί η κλήση που στέλνει ο πελάτης προς τον εξυπηρέτη. Κάθε φορά που ο πελάτης στέλνει μια κλήση ενεργοποιεί μία προθεσμία απόκρισης (timeout). Αν η προθεσμία λήξει πριν ληφθεί το αντίστοιχο μήνυμα απόκρισης, ο πελάτης επαναμεταδίδει το μήνυμα κλήσης (request re-invocation). Εάν το μήνυμα έχει πραγματικά χαθεί, ο εξυπηρέτης δε θα είναι σε θέση να ξεχωρίσει το αρχικό από το επαναμεταδιδόμενο και το σύστημα θα συνεχίσει τη λειτουργία του χωρίς πρόβλημα. Βέβαια, εάν χάνονται πολλά μηνύματα τότε ο πελάτης μπορεί να θεωρήσει ότι ο εξυπηρέτης είναι εκτός λειτουργίας, οπότε επιστρέφουμε στην πρώτη περίπτωση.
3. Ένα άλλο σφάλμα που μπορεί να συμβεί είναι να έχουμε απωλεσθέντα μηνύματα απόκρισης. Αυτό θα πει πως ο εξυπηρέτης λαμβάνει κανονικά το μήνυμα κλήσης από τον πελάτη, διεκπεραιώνει την κλήση, στέλνει το μήνυμα απόκρισης στον πελάτη, αλλά αυτό χάνεται. Ο πελάτης βέβαια, αν δε λάβει απάντηση μέσα στην καθορισμένη προθεσμία, στέλνει ξανά το μήνυμα κλήσης (request re-invocation). Το πρόβλημα σε

αυτήν την περίπτωση είναι ότι ο πελάτης δεν είναι σε θέση να γνωρίζει τους λόγους για τους οποίους δεν έλαβε απόκριση, δηλαδή δεν μπορεί να ξέρει για παράδειγμα αν χάθηκε η κλήση, αν χάθηκε η απόκριση, αν ο εξυπηρέτης είναι αργός κτλ.

4. Ένα άλλο πιθανό σφάλμα είναι η κατάρρευση του εξυπηρέτη. Σε αυτό το σφάλμα υπάρχουν δύο υποπεριπτώσεις. Στην πρώτη υποπερίπτωση η σειρά των γεγονότων είναι η εξής: η κλήση φτάνει στον εξυπηρέτη, διεκπεραιώνεται, αλλά ο εξυπηρέτης καταρρέει πριν προλάβει να στείλει την απόκριση. Στην δεύτερη υποπερίπτωση και πάλι η κλήση φτάνει στον εξυπηρέτη, αλλά ο εξυπηρέτης καταρρέει αμέσως μετά. (Η περίπτωση 4 απεικονίζεται στο Σχήμα 3.)

5. Ένα ακόμη ενδεχόμενο σφάλμα είναι να καταρρεύσει ο πελάτης. Να στείλει δηλαδή μια αίτηση στον εξυπηρέτη και να τεθεί εκτός λειτουργίας πριν λάβει την απόκριση από τον εξυπηρέτη.

Όλα τα παραπάνω σφάλματα που αναφέραμε μπορούν να αντιμετωπιστούν και να διορθωθούν. Επίσης όλα τα παραπάνω σφάλματα είναι δυνατόν να επανεμφανιστούν μετά την επανόρθωσή τους. Υπάρχουν όμως και σφάλματα που μπορεί να συμβούν σε ένα κατανεμημένο σύστημα, αλλά να αργήσουν πολύ να επανεμφανιστούν μετά την επανόρθωσή τους. Τέτοια σφάλματα είναι εξ ορισμού σπάνια γεγονότα και μπορεί να είναι για παράδειγμα ανεπαρκής μνήμη, διακοπή παροχής ενέργειας και άλλα.



Σχήμα 3. (α) Κανονική Περίπτωση. (β) Ο εξυπηρέτης καταρρέει μετά την εκτέλεση. (γ) Ο εξυπηρέτης καταρρέει πριν την εκτέλεση.

1.3 Πλεονασματική Επεξεργασία Αντικειμένων

Όπως ήδη αναφέραμε, ο όρος ομάδα αντικειμένων (object group) αναφέρεται στο σύνολο των πλεονασματικών αντιγράφων ενός αντικειμένου (όπου το αντικείμενο αυτό

μπορεί να είναι είτε εξυπηρέτης είτε πελάτης). Ένα από τα αντίγραφα του αντικειμένου το ονομάζουμε *κύριο (primary replica)* και τα υπόλοιπα τα ονομάζουμε *εφεδρικά (backup replicas).* Ο λόγος για τον οποίο δημιουργούμε πολλά αντίγραφα είναι ο εξής: αν σε ένα αντικείμενο διακοπεί η λειτουργία κάποιου αντιγράφου (είτε κύριου είτε εφεδρικού), θέλουμε να υπάρχουν επιπλέον αντίγραφα έτσι ώστε ένα από αυτά να το αντικαταστήσουν συνεχίζοντας τις λειτουργίες του. Η κατάσταση αυτή θα κρατήσει μέχρι να γίνει επαναφορά (recovering) του αντιγράφου στο οποίο συνέβη το σφάλμα. Τότε λοιπόν, το αντίγραφο αυτό θα βρεθεί ξανά στην κανονική κατάσταση λειτουργίας του, είτε ως κύριο είτε ως εφεδρικό. Οι καταστάσεις στις οποίες μπορεί να βρεθεί ένα αντίγραφο, είτε αυτό είναι κύριο είτε εφεδρικό, εξαρτάται από το είδος πλεονασματικής επεξεργασίας που χρησιμοποιείται. Τα είδη πλεονασματικής επεξεργασίας αναλύονται παρακάτω.

Σε κάποια συστήματα υπάρχει η δυνατότητα να δημιουργούνται νέα αντίγραφα ή να διαγράφονται ήδη υπάρχοντα αντίγραφα σε ένα αντικείμενο. Το αρχικό πλήθος των αντιγράφων που δημιουργούνται σε ένα αντικείμενο ονομάζεται *αρχικός αριθμός αντιγράφων (initial number of replicas).* Το ελάχιστο πλήθος αντιγράφων που πρέπει να υπάρχουν σε ένα αντικείμενο έτσι ώστε αυτό να είναι επαρκώς προστατευμένο από τα λάθη ονομάζεται *ελάχιστος αριθμός αντιγράφων (minimum number of replicas).*

Υπάρχουν τρία *είδη πλεονασματικής επεξεργασίας* (*replication styles*) και είναι τα ακόλουθα:

1.3.1 Ενεργητική πλεονασματική επεξεργασία (Active Replication)

Σε αυτήν την περίπτωση δεν υπάρχει διαχωρισμός των αντιγράφων σε κύρια και εφεδρικά γιατί όλα τα αντίγραφα του αντικειμένου εκτελούν όλες τις λειτουργίες με την ίδια σειρά, ανεξάρτητα όμως το ένα από το άλλο. Βέβαια δε δουλεύουν όλα τα αντίγραφα με την ίδια ταχύτητα. Παρόλα αυτά όμως μετά το τέλος μιας λειτουργίας, όλα τα αντίγραφα που έφεραν σε πέρας την λειτουργία αυτή έχουν την ίδια κατάσταση και το γεγονός αυτό για την περίπτωση της ενεργητικής πλεονασματικής επεξεργασίας χαρακτηρίζεται από τον όρο *ισχυρή συνέπεια ομάδας αντικειμένων* (*strong replica consistency*). Με αυτόν τον τρόπο εάν συμβεί σφάλμα σε ένα αντίγραφο, αυτό θα

sistency). Με αυτόν τον τρόπο εάν συμβεί σφάλμα σε ένα αντίγραφο, αυτό θα καλυφθεί από τα υπόλοιπα αντίγραφα, τα οποία θα συνεχίσουν κανονικά την εκτέλεση της λειτουργίας χωρίς να επηρεαστούν από το σφάλμα, δηλαδή χωρίς να περιμένουν να ανιχνευτεί το σφάλμα και να διορθωθεί.

Στην περίπτωση της ενεργητικής πλεονασματικής επεξεργασίας, θα μπορούσε ο πελάτης να λαμβάνει πολλαπλές απαντήσεις για ένα αίτημά του ή αντίστοιχα ο εξυπηρέτης να λαμβάνει το ίδιο αίτημα για εξυπηρέτηση πολλές φορές. Αυτό θα γινόταν επειδή όπως είπαμε όλα τα αντίγραφα του αντικειμένου δουλεύουν παράλληλα και ανταποκρίνονται σε κάθε αίτημα που φτάνει στο αντικείμενο (είτε αυτό είναι πελάτης είτε εξυπηρέτης). Όμως τα διπλά αιτήματα καθώς και οι διπλές απαντήσεις πρέπει να ανιχνεύονται και να καταστέλλονται και έτσι να αποστέλλεται κάθε φορά μία μόνο αίτηση ή μία μόνο απάντηση στο εκάστοτε αντικείμενο.

1.3.2 Παθητική πλεονασματική επεξεργασία: Θερμή πλεονασματική επεξεργασία (Warm Passive Replication)

Σε αυτό το μοντέλο μόνο ένα από τα αντίγραφα, το χαρακτηριζόμενο ως κύριο, διεκπεραιώνει όλες τις λειτουργίες εφόσον βρίσκεται στην κανονική κατάσταση λειτουργίας του. Η κατάσταση του κύριου αντιγράφου και η σειρά των λειτουργιών που περατώνει αποθηκεύονται σε μια ουρά καταγραφής (*message log*). Τα υπόλοιπα αντίγραφα συγχρονίζονται με το κύριο σε τακτά χρονικά διαστήματα. Έτσι όταν θα συμβεί σφάλμα στο κύριο αντίγραφο, τότε κάποιο από τα εφεδρικά θα το αντικαταστήσει.

Οι δυνατές καταστάσεις στις οποίες μπορεί να βρεθεί ένα αντίγραφο στην περίπτωση της θερμής παθητικής πλεονασματικής επεξεργασίας είναι: α) κανονική κατάσταση (normal), β) κατάσταση σφάλματος (fault), γ) κατάσταση επαναφοράς (recovering) και δ) κατάσταση μεταφοράς/αντιγραφής κατάστασης (state transferring). Οι μεταβάσεις μεταξύ των καταστάσεων είναι διαφορετικές για την περίπτωση του κύριου αντιγράφου και των εφεδρικών αντιγράφων και φαίνονται στο Σχήμα 5 (κεφάλαιο 5).

1.3.3 Παθητική πλεονασματική επεξεργασία: Ψυχρή πλεονασματική επεξεργασία (Cold Passive Replication)

Και σε αυτήν την περίπτωση είναι το κύριο αντίγραφο μόνο που διεκπεραιώνει όλες τις λειτουργίες και η κατάσταση του κύριου αντιγράφου αποθηκεύεται και πάλι περιοδικά σε μια ουρά καταγραφής (message log). Τα υπόλοιπα αντίγραφα μένουν αδρανή για όσο διάστημα το κύριο αντίγραφο λειτουργεί κανονικά. Όταν όμως συμβεί σφάλμα, ένα από τα εφεδρικά αντίγραφα φορτώνεται στη μνήμη και ανακτά την κατάσταση του κύριου αντιγράφου από την ουρά καταγραφής (message log) και έτσι αναλαμβάνει αυτό όλες τις λειτουργίες μέχρι να επανορθωθεί το κύριο αντίγραφο. Οι δυνατές καταστάσεις στις οποίες μπορεί να βρεθεί ένα αντίγραφο στην περίπτωση της ψυχρής παθητικής πλεονασματικής επεξεργασίας είναι ίδιες με την περίπτωση της θερμής παθητικής πλεονασματικής επεξεργασίας.

1.4 Ανίχνευση Σφαλμάτων

Για την ανίχνευση σφαλμάτων θεωρούμε ότι υπάρχει ένας τοπικός ανιχνευτής για κάθε αντικείμενο και ένας γενικός ανιχνευτής που παρακολουθεί τους τοπικούς ανιχνευτές. Στο σύστημα που αναπτύξαμε κάθε αντικείμενο ελέγχεται περιοδικά σύμφωνα με ένα προκαθορισμένο χρονικό διάστημα και αν βρεθεί κάποιο σφάλμα ακολουθούνται οι διαδικασίες που αναφέρθηκαν παραπάνω.

ΚΕΦΑΛΑΙΟ 2ο

Σύντομη Αναφορά στην Σχετική Βιβλιογραφία

Τα αναλυτικά μοντέλα απόδοσης (analytic performance models) για λογισμικό πλεονασματικής επεξεργασίας έχουν επικεντρωθεί στην αξιολόγηση πλεονασματικής επεξεργασίας διεργασιών (process-based replication) κι όχι αντικειμένων (object-based replication schemes). Στο άρθρο [4], οι συγγραφείς εισάγουν ένα αναλυτικό μοντέλο που το ονομάζουν Fault-Tolerant Layered Queuing Network (FTLQN). Αυτό το μοντέλο προβλέπει αποτελεσματικά την εκτελεσιμότητα κατανεμημένων συστημάτων με ανοχή σε λάθη που ακολουθούν το μοντέλο εξυπηρετούμενου-εξυπηρέτη. Επίσης, το παραπάνω μοντέλο είναι κατάλληλο για συστήματα τα οποία χρησιμοποιούν εναλλακτικούς εξυπηρετητές και εναλλακτική δρομολόγηση των αιτημάτων ώστε να καλύψουν τα σφάλματα.

Στο άρθρο [24] έχει αναφερθεί αξιολόγηση απόδοσης που να βασίζεται στην προσομοίωση. Τα σχετικά πειράματα που έχουν διεξαχθεί, επικεντρώνονται ξεχωριστά γύρω από: α) την τεχνική καταγραφής μηνυμάτων (message logging) των διεργασιών που εξυπηρετούνται, β) τα σημεία καταγραφής κατάστασης (checkpointing) και γ) τα σχήματα επαναφοράς κατάστασης (recovery schemes). Οι υλοποιήσεις σε λογισμικά προσομοίωσης των προαναφερθέντων περιπτώσεων όμως, δεν μπορούν να χρησιμοποιηθούν για την αξιολόγηση σχημάτων πλεονασματικής επεξεργασίας (replication schemes) που αποτελούνται από πιθανώς διαφορετικές πολιτικές για τα αντικείμενα που απαρτίζουν το σύστημα. Επιπλέον, δεν γίνεται διαχωρισμός μεταξύ της αποτελεσματικότητας της ανοχής σε λάθη και της απόδοσης της ανοχής σε λάθη, κάτι που όμως υπάρχει στην προσέγγιση που η παρούσα εργασία παρουσιάζει.

Τέλος, αξίζει να αναφερθεί η δημοσίευση του άρθρου [15], στο οποίο οι συγγραφείς προτείνουν έναν υβριδικό αλγόριθμο μαθηματικού προγραμματισμού και αναλυτικής

αξιολόγησης, για ένα πρόβλημα μελέτης παραγόντων αντιστάθμισης (trade-off problem) που δεν έχει άμεση σχέση με ανοχή σε λάθη: καθορίζονται τα επίπεδα πλεονασματικής επεξεργασίας των διεργασιών ή τα επίπεδα πολυνηματισμού, με τέτοιο τρόπο, όμως, ώστε να αποφευχθούν μη απαραίτητες καθυστερήσεις αναμονής στις ουρές των αποστολέων των αιτήσεων ή ώστε να αποφευχθεί μη απαραίτητη υψηλή κατανάλωση μνήμης.

ΚΕΦΑΛΑΙΟ 3ο

Αξιολόγηση Σχημάτων Πλεονασματικής Επεξεργασίας Αντικειμένων σε Εφαρμογές Αξιόπιστης Εξυπηρέτησης

3.1 Προεπισκόπηση

Εφαρμογές αξιόπιστης εξυπηρέτησης με αντικειμενοστραφή σχεδίαση (object-based dependable server applications) κάνουν χρήση σχημάτων ανοχής σε σφάλματα (fault tolerance schemes), που πιθανώς όμως αποτελούνται από διαφορετικές πολιτικές πλεονασματικής επεξεργασίας (replication policies) για τα αντικείμενα που τις συνθέτουν. Τέτοια σχήματα τα καλούμε *σύνθετα σχήματα πλεονασματικής επεξεργασίας (composite replication schemes)*. Η εργασία αυτή (καθώς και το άρθρο [30]) παρουσιάζει μια προσέγγιση εκτίμησης βασισμένη σε προσομοίωση (simulation-based evaluation approach).

Σε σύγκριση με άλλες προσεγγίσεις εκτίμησης :
(α) η προσέγγιση που δε στηρίζεται στη χρήση της γνωστής προσομοίωσης αξιοπιστίας συστήματος (system's reliability simulation), αλλά σε μια υβριδική προσομοίωση *αξιοπιστίας* ***και*** *κίνησης μηνυμάτων* (hybrid reliability and system's traffic simulation) και
(β) κάνουμε μια ξεκάθαρη και σαφή διάκριση μεταξύ των μετρικών που χρησιμοποιούνται για τους χρόνους απόκρισης των λειτουργιών που επηρεάζονται από σφάλματα (fault-affected service response times) και των χρόνων απόκρισης των λειτουργιών που δεν έχουν επηρεαστεί από σφάλματα (fault-unaffected service response times).

Το πρώτο χαρακτηριστικό (το (α)) μας επιτρέπει να λαμβάνουμε υπόψη διαφορετικά θέματα (όπως ανοχή σε λάθη, εξισορρόπηση υπολογιστικού φόρτου και πολυνηματισμό) με έναν κοινό και συνδυασμένο τρόπο. Το δεύτερο χαρακτηριστικό καθιστά την προτεινόμενη προσέγγιση, κατάλληλη για τη σχεδίαση συστημάτων, όπου η πρόθεση είναι ο προσδιορισμός είτε των βέλτιστων τιμών των παραμέτρων πλεονασματικής επεξεργασίας (*optimal* replication properties), είτε των εγγυήσεων χρόνων απόκρισης για τους χρήστες του συστήματος.

Στο Κεφάλαιο 7 παρουσιάζουμε αποτελέσματα που αφορούν την ανάλυση μιας συγκεκριμένης περίπτωσης συστήματος. Τα πειράματα που έγιναν βασίστηκαν πάνω σε διαφορετικές υποθέσεις σχετικά με το τι γίνεται όταν συμβαίνουν σφάλματα σε έναν εξυπηρέτη (loss scenarios). Τα αποτελέσματα αυτά παρέχουν γνώση για το σχεδιασμό σύνθετων στρατηγικών αναμονής απόκρισης - επανάκλησης (composite request-retry schemes) με κατάλληλους χρόνους προθεσμίας απόκρισης (appropriate request time-outs).

3.2 Η μεθοδολογία εκτίμησης

Στις εφαρμογές αξιόπιστης εξυπηρέτησης η ανοχή σε λάθη βασίζεται πάνω στην χρήση πλεονασματικής επεξεργασίας ώστε να ελαχιστοποιείται η απώλεια υπολογισμών στην παρουσία μη επαναλαμβανόμενων λαθών. Συνηθισμένες αιτίες λαθών που δεν επαναλαμβάνονται μετά την επανάκτηση (recovery) είναι: ανεπαρκής μνήμη, σφάλματα υλικού, διακοπές ρεύματος και σφάλματα δικτύου (βλ. Κεφάλαιο 1) κι ο μη ντετερμινισμός που εισάγεται είτε από τη χρήση κατανεμημένων χρονικών μηχανισμών είτε από την χρήση πολυνηματισμού.

Τα σχήματα ανοχής λαθών για εφαρμογές αντικειμενοστραφούς σχεδίασης είναι πιθανό να αποτελούνται από διαφορετικές πολιτικές πλεονασματικής επεξεργασίας για τα αντικείμενα που συνθέτουν την εφαρμογή (σύνθετα σχήματα πλεονασματικής επεξεργασίας). Σύμφωνα με τις πρόσφατα δημοσιευθείσες προδιαγραφές του OMG FT-CORBA ([21]), το είδος της πλεονασματικής επεξεργασίας που ανατίθεται σε ένα αντικείμενο μπορεί να είναι ενεργητική, θερμή παθητική ή ψυχρή παθητική πλεονασματική επεξεργασία (βλ. Κεφάλαιο 1) και παραμετροποιείται από ένα σύνολο

ιδιοτήτων συμπεριφοράς (πλήθος αντιγράφων, σημεία καταγραφής κατάστασης, διαστήματα αντιγραφής κατάστασης, διάστημα αναμονής απόκρισης-επανάληψης, κτλ.) που κάθε φορά είναι κατάλληλο σύμφωνα με το εφαρμοζόμενο περιβάλλον ανίχνευσης λαθών.

Στις εφαρμογές αξιόπιστης εξυπηρέτησης όπου η ανοχή σε λάθη επιτυγχάνεται μέσω σύνθετων σχημάτων πλεονασματικής επεξεργασίας, η ποιότητα εξυπηρέτησης κυριαρχείται από χαρακτηριστικούς παράγοντες αντιστάθμισης της αποτελεσματικότητας της ανοχής σε λάθη κι της απόδοσης:

- Υπερβολικά συχνά σημεία καταγραφής κατάστασης και διαστήματα αντιγραφής κατάστασης μεταξύ αντικειμένων έχουν σαν αποτέλεσμα την υποβάθμιση της απόδοσης, ενώ ανεπαρκής τοποθέτηση σημείων καταγραφής και αραιά διαστήματα αντιγραφής κατάστασης προκαλούν ακριβή επανάκτηση.
- Υπερβολικά συχνά διαστήματα αναμονής απόκρισης-επανάκλησης στα αντικείμενα προκαλούν υψηλά επιπλέον κόστη και δεν βελτιώνουν την αποτελεσματικότητα της ανοχής σε λάθη.

Η ποιότητα εξυπηρέτησης επίσης εξαρτάται κι από δομικές εξαρτήσεις που επιβάλλονται από την ακολουθία ανταλλαγής των μηνυμάτων.

Ένα σύνθετο σχήμα πλεονασματικής επεξεργασίας μπορεί να συνδυαστεί με κάποια συγκεκριμένη στρατηγική ανάθεσης των αιτημάτων στους εξυπηρετητές ώστε να επιτευχθεί η παροχή ενός προσυμφωνημένου επιπέδου ποιότητας εξυπηρέτησης. Τα σημεία καταγραφής και αντιγραφής κατάστασης επηρεάζουν την αποστολή των αιτήσεων και γι' αυτόν τον λόγο επηρεάζουν και την απόδοση του εφαρμοζόμενου φόρτου εργασίας.

Μέσα σε αυτό το υπολογιστικό περιβάλλον, όταν μια εφαρμογή πρέπει να συμμορφώνεται σε πιθανώς προσυμφωνημένες εγγυήσεις ποιότητας εξυπηρέτησης όσον αφορά τους χρόνους απόκρισης, οι μέθοδοι αξιολόγησης της αξιοπιστίας που έχουν δημοσιευθεί είναι ανεπαρκείς [10].

Η εργασία αυτή παρουσιάζει μια υβριδική προσομοίωση αξιοπιστίας *και* κυκλοφορίας συστήματος: γεγονότα που συμβαίνουν σε διαφορετικές χρονικές κλίμακες – όπως πχ. αιφνίδια λάθη, τα οποία εξ ορισμού είναι σπάνια γεγονότα, κι αφίξεις αιτημάτων που

είναι πολύ συχνά γεγονότα – προσομοιώνονται μαζί στο ίδιο πείραμα. Το γεγονός αυτό επιτρέπει να λαμβάνουμε υπόψη με ένα συνδυασμένο τρόπο περίπλοκες αλληλεπιδράσεις που διαφορετικά θα αποδίδονταν ξεχωριστά σε διάφορες σχεδιαστικές επιλογές (ανοχή σε λάθη, εξισορρόπηση φόρτου, πολυνηματισμός).

Η προσέγγιση αξιολόγησης που παρουσιάζεται σε αυτήν την εργασία, επίσης, διαφέρει συγκρινόμενη με άλλες όσον αφορά τα μέτρα αξιολόγησης. Κάνουμε μία διάκριση μεταξύ των μέτρων που χρησιμοποιούνται για τους χρόνους απόκρισης των αιτημάτων που επηρεάζονται από λάθη κι εκείνων που δεν επηρεάζονται από λάθη. Η πρώτη κατηγορία χρόνων απόκρισης ποσοτικοποιεί την επιτυγχανόμενη αποτελεσματικότητα της ανοχής στα λάθη, ενώ η δεύτερη αναφέρεται στη μεγάλη πλειοψηφία των αιτημάτων κι άρα χαρακτηρίζει την προκύπτουσα απόδοση της ανοχής σε λάθη.

Η ποιότητα εξυπηρέτησης και οι σχετιζόμενες εγγυήσεις αφορούν όμως και τις δυο κατηγορίες αιτημάτων. Αυτή η εργασία δεν αποσκοπεί απλά στο να παρουσιάσει μια ακόμη ικανοποιητική προσέγγιση αξιολόγησης. Η χωριστή ποσοτικοποίηση της απόδοσης και της αποτελεσματικότητας της ανοχής σε λάθη, που παρουσιάζει η εργασία, παρέχει μια διαίσθηση για τους πιο σημαντικούς παράγοντες αντιστάθμισης μεταξύ αποτελεσματικότητας και απόδοσης και για τις περίπλοκες αλληλεπιδράσεις που προκαλούνται από τις δομικές εξαρτήσεις του συστήματος. Η προσέγγιση αξιολόγησης, που παρουσιάζει αυτή η εργασία, μπορεί να αποτελέσει πλέον βασικό εργαλείο συστηματικών μεθόδων σχεδίασης συστημάτων παροχής, όπως πχ. η μέθοδος που παρουσιάζεται στο [11].

Στην συνέχεια της εργασίας περιγράφεται η προσέγγιση αξιολόγησης και η λειτουργικότητα του πρωτότυπου λογισμικού προσομοίωσης, όπου εφαρμόστηκε η παρουσιαζόμενη προσέγγιση αποτίμησης. Παρέχονται αποτελέσματα αναφοράς από μια μελέτη περίπτωσης (case study) που βασίζεται σε διαφορετικές υποθέσεις για το τι συμβαίνει όταν σε ένα αντικείμενο παρουσιάζεται σφάλμα (loss scenarios) ή αλλιώς όπως θα τα αποκαλούμε *διαφορετικά σενάρια απώλειας κλήσεων*. Τα αποτελέσματα παρέχουν μία εικόνα για τον σχεδιασμό σύνθετων σχημάτων αναμονής απόκρισης-επανάκλησης συνδυαζόμενα με κατάλληλα διαστήματα αποστολής αιτημάτων.

Το επόμενο κεφάλαιο καθορίζει λεπτομερώς το υποτιθέμενο υπολογιστικό μοντέλο και τα προσομοιωμένα μοντέλα λαθών αντικειμένων. Το κεφάλαιο 5 παρουσιάζει την λειτουργικότητα που παρέχεται από το ανεπτυγμένο πρωτότυπο λογισμικό προσομοίωσης όσον αφορά την συνδυασμένη αξιολόγηση πολυνηματισμού, εξισορρόπησης φόρτου, πλεονασματικής επεξεργασίας και πολιτικών απόκρισης-επανάκλησης, κάτω από διαφορετικές προδιαγραφές ανίχνευσης λαθών. Το κεφάλαιο 6 επικεντρώνεται στην παρουσίαση της προσέγγισης αξιολόγησης. Το κεφάλαιο 7 παρουσιάζει το μοντέλο της μελέτης περίπτωσης και συνοψίζει τα αποτελέσματα των πειραμάτων. Τελικά, καταλήγουμε με μια συζήτηση πάνω στις προοπτικές της παρουσιαζόμενης προσέγγισης αποτίμησης και των τρόπων χρήσης του πρωτότυπου λογισμικού προσομοίωσης.

ΚΕΦΑΛΑΙΟ 4ο

Υπολογιστικό Μοντέλο και Μοντέλα Λαθών

4.1 Το Υπολογιστικό Μοντέλο

Το υπολογιστικό μοντέλο, πάνω στο οποίο βασίζεται αυτή η εργασία, είναι το OMG Core Object Model – όπως τεκμηριώνεται στο [22] και στο Κεφάλαιο 2 – εμπλουτισμένο από τις υποθέσεις που υιοθετεί το πρότυπο OMG FT-CORBA ([23]). Στην συνέχεια, σκιαγραφούνται οι υποθέσεις που είναι απαραίτητες για την παρουσίαση της προσέγγισης αποτίμησης:

- Ένα αντικείμενο μπορεί να μοντελοποιήσει οποιοδήποτε είδος οντότητας ή έννοιας και χαρακτηρίζεται από την δική του διακριτή ταυτότητα, η οποία παραμένει αμετάβλητη και μόνιμη, και διαρκεί για όσο διάστημα το αντικείμενο υπάρχει κι είναι ανεξάρτητη από τις ιδιότητες του αντικειμένου ή την συμπεριφορά του. *Κάθε αντικείμενο είτε κατέχει ή είτε δεν κατέχει μια κατάσταση* (state) και ένα σύνολο μεθόδων (συναρτήσεων - διαδικασιών).
- Κάθε μέθοδος έχει μια υπογραφή η οποία αποτελείται από το όνομα της μεθόδου, το σύνολο των παραμέτρων (εισόδους στη μέθοδο) κι από τα αποτελέσματα (εξόδους της μεθόδου). Μια μέθοδος περιγράφει μια ενέργεια ή λειτουργία που μπορεί να εφαρμοστεί πάνω στις παραμέτρους της. Μια κλήση, ή αλλιώς ενεργοποίηση, μεθόδου (method invocation), που

ονομάζεται και αίτηση ή αίτημα (request), υποδηλώνει μια μέθοδο, αναθέτει κάποιες παραμέτρους εκ μέρους κάποιου αιτούμενου (client) και προκαλεί την ενεργοποίηση της και την επιστροφή αποτελεσμάτων. Οι μέθοδοι ενός αντικειμένου αποτελούν τον μόνο τρόπο μέσω του οποίου μπορεί να αλλάξει η κατάστασή του. Στην εργασία αυτή θεωρείται μόνο η χρήση σύγχρονης κλήσης μεθόδων, που αποτελεί τον πιο διαδεδομένο μηχανισμό επικοινωνίας μεταξύ διαδικασιών.

Σε μια σύγχρονη κλήση μεθόδου (synchronous method invocation) το αντικείμενο που αιτείται την εξυπηρέτηση (δηλαδή την εκτέλεση της ενεργοποιημένης μεθόδου) σταματάει την δική του εκτέλεση και περιμένει την μέθοδο που ενεργοποίησε να τερματίσει και να επιστρέψει τα αποτελέσματα. Μόλις λάβει την απάντηση, ο αιτούμενος συνεχίζει την εκτέλεσή του από το σημείο διακοπής.

- Υιοθετείται η *το-πολύ-μια-φορά κλήση* (at-most-once invocation) σημασιολογία του OMG Core Object Model, που σημαίνει ότι κάθε εξυπηρέτηση αίτησης εκτελείται το πολύ μια φορά. Διπλές, ή και πολλαπλές, κλήσεις που μπορεί να υπάρξουν εξαιτίας της πλεονασματικής επεξεργασίας (replication) εντοπίζονται και καταστέλλονται.
- Στο OMG FT-CORBA, η ανοχή σε σφάλματα βασίζεται στη δημιουργία και διαχείριση πολλαπλών αντιγράφων αντικειμένου ως μια μοναδική ομάδα αντικειμένων (object group). Τα αντικείμενα που είναι εξυπηρετούμενοι ενεργοποιούν μεθόδους στην ομάδα αντικειμένων εξυπηρέτησης και ένα ή περισσότερα μέλη της ομάδας εξυπηρέτησης εκτελούν τις μεθόδους και επιστρέφουν τις απαντήσεις τους στους εξυπηρετούμενους, όπως θα έκανε κι ένα συμβατικό αντικείμενο.
- Δεν λαμβάνεται υπόψη η τοπολογία του δικτύου ή τα πρωτοκολλά πάνω στα οποία βασίζεται η επικοινωνία μεταξύ των διαδικασιών και των υπολογιστικών κέντρων (hosts). Θεωρείται μόνο ότι παρέχουν *αξιόπιστη πλήρως διατεταγμένη παράδοση των αιτήσεων* (reliable totally ordered delivery of requests) στα αντίγραφα κάθε αντικειμένου.
- Στο OMG FT-CORBA ορίζεται η *ισχυρή συνέπεια αντιγράφων* (strong replica consistency), σύμφωνα με την οποία, ακόμη και στην παρουσία σφαλμάτων, καθώς τα μέλη μιας ομάδας αντικειμένων εκτελούν την ακολουθία των μεθόδων, η συμπεριφορά της ομάδας πρέπει να είναι λογικά

ισοδύναμη με αυτήν ενός μοναδικού, ελεύθερου από σφάλματα αντικειμένου που επεξεργάζεται την ακολουθία κλήσεων μεθόδων. Για κάθε αντικείμενο ξεχωριστά η έννοια της ισχυρής συνέπειας ομάδας αντικειμένων διατηρεί το κατάλληλο, αντίστοιχο περιεχόμενο που εξαρτάται από *είδος πλεονασματικής αντιγραφής* (replication style, βλ. Κεφάλαια 1 και 5).

- Οι αξιόπιστες εφαρμογές εξυπηρέτησης μπορούν να επωφεληθούν από το πολυνηματικό Object Request Brokers ([27]), που ορίζει ότι η σειρά της νηματικής εκτέλεσης των διακινούμενων μεθόδων πρέπει να καθορίζεται ντετερμινιστικά είτε για όλες τις κλήσεις ([18]), είτε μόνο για εκείνες που εξαρτώνται από άλλες ([1]). Υιοθετείται το MT-Domain αφαιρετικό μοντέλο ([18]) για την αναφορά σε οποιοδήποτε αντικείμενο εξυπηρέτησης ή εξυπηρετούμενου που υποστηρίζει πολλαπλά νήματα – και που ίσως έχει πρόσβαση σε διαμοιραζόμενα δεδομένα – και που μπορεί να περιέχει ένα ή περισσότερα αντικείμενα.
- Θεωρείται πως το υποκείμενο λειτουργικό σύστημα επιδεικνύει ντετερμινιστική συμπεριφορά και επιπλέον, πως η διαχείριση των όποιων κλήσεων συστήματος (αν συμβαίνουν) γίνεται από έναν κατάλληλο μηχανισμό ο οποίος δεν εισάγει μη ντετερμινισμό.
- Η εξισορρόπηση φόρτου (load balancing) είναι ένα ζήτημα που δεν καλύπτεται από τις μέχρι τώρα δραστηριότητες τυποποίησης. Δεν γίνεται διαχωρισμός μεταξύ συγκεντρωτικού ή κατανεμημένου φόρτου (centralized or distributed load balancing). Οι πληροφορίες περί φόρτου, αν είναι δυνατό, παρέχονται από την υποκείμενη υποδομή (*middleware* infrastructure) του συστήματος, η οποία αναθέτει αιτήματα στα διαθέσιμα αντικείμενα εξυπηρέτησης με τρόπο που είναι διαφανής και επιδεικνύει ανοχή σε σφάλματα (όπως στο [28] και στο [14]).

4.2 Μοντέλα Σφαλμάτων

Όπως έχει ήδη αναφερθεί στην (εισαγωγή), η εργασία ασχολείται με σφάλματα που δεν επαναλαμβάνονται μετά την αντιμετώπισή τους και που καλούνται *προσωρινά σφάλματα αντικειμένων* (transient object faults). Τα αντικείμενα των εφαρμογών

συμμορφώνονται στο *μοντέλο σφάλμα-παύση* (fail-stop model) ([26]), που σημαίνει ότι όταν συμβαίνουν σφάλματα τα αντικείμενα σταματούν πλήρως οποιαδήποτε λειτουργία, καταρρέουν εντελώς, χωρίς να στέλνουν ψεύτικα (spurious) μηνύματα. Αυτή η εργασία *δεν* ασχολείται με διαδιδόμενα σφάλματα (commission faults), όπως είναι για παράδειγμα τα Βυζαντινά σφάλματα. Δεν ασχολείται δηλαδή, με περιπτώσεις όπου μετά το σφάλμα το αντικείμενο συνεχίζει την λειτουργία του παρέχοντας όμως λανθασμένα αποτελέσματα, διαδίδοντας πιθανώς το σφάλμα σε άλλα αντικείμενα.

Η εργασία λαμβάνει υπόψη σφάλματα δικτύου (network faults), τα οποία όμως όταν συμβαίνουν ο πελάτης δεν μπορεί να τα ανιχνεύσει. Επίσης, όταν συμβαίνουν σφάλματα δικτύου, ο αιτούμενος δεν λαμβάνει απάντηση, την οποία όμως άδικα περιμένει. Όπως και στο OMG FT-CORBA, στην εργασία αποκλείονται σφάλματα δικτύου τα οποία χωρίζουν τους εξυπηρετητές (hosts, servers) του συστήματος σε δυο ή περισσότερα σύνολα (network-partitioning faults).

Γενικά, *όταν σε ένα αντικείμενο εξυπηρετητή συμβαίνει σφάλμα, οι αιτήσεις που βρίσκονται ήδη σε αναμονή στην ουρά δεν χάνονται κι όλα τα αιτήματα που καταφθάνουν ενόσω το αντικείμενο δεν είναι διαθέσιμο ή/και λειτουργικό τοποθετούνται στο τέλος της ουράς*. Για τα αντικείμενα παθητικής πλεονασματικής επεξεργασίας θεωρούμε κι επιπλέον σενάρια απώλειας κλήσεων (loss behaviors):

(α) Όταν σε ένα αντικείμενο συμβαίνει σφάλμα, τα ήδη υπάρχοντα αιτήματα στην ουρά δεν χάνονται, αλλά τα νέα αιτήματα που καταφθάνουν ενόσω το αντικείμενο δεν είναι διαθέσιμο ή/και λειτουργικό χάνονται.

(β) Όταν σε ένα αντικείμενο συμβαίνει σφάλμα, από τα ήδη υπάρχοντα αιτήματα στην ουρά, μόνο εκείνο που βρίσκεται σε εξυπηρέτηση (αν υπάρχει τέτοιο) χάνεται, αλλά και τα νέα αιτήματα που καταφθάνουν ενόσω το αντικείμενο δεν είναι διαθέσιμο ή/και λειτουργικό επίσης χάνονται.

(γ) Όταν σε ένα αντικείμενο συμβαίνει σφάλμα, από τα ήδη υπάρχοντα αιτήματα στην ουρά μόνο εκείνο που βρίσκεται σε εξυπηρέτηση (αν υπάρχει τέτοιο) χάνεται, αλλά τα νέα αιτήματα που καταφθάνουν ενόσω το αντικείμενο δεν είναι διαθέσιμο ή/και λειτουργικό τοποθετούνται στο τέλος της ουράς.

(δ) Όταν σε ένα αντικείμενο συμβαίνει σφάλμα, τα ήδη υπάρχοντα αιτήματα στην ουρά χάνονται, αλλά και τα νέα αιτήματα που καταφθάνουν ενώ το αντικείμενο δεν είναι διαθέσιμο ή/και λειτουργικό επίσης χάνονται.

(ε) Όταν σε ένα αντικείμενο συμβαίνει σφάλμα, τα ήδη υπάρχοντα αιτήματα στην ουρά χάνονται, αλλά τα νέα αιτήματα που καταφθάνουν ενώ το αντικείμενο δεν είναι διαθέσιμο ή/και λειτουργικό τοποθετούνται στο τέλος της ουράς.

Στα πλαίσια αυτής της εργασίας αναπτύχθηκε ένας πρωτότυπος και με δυνατότητες επέκτασής αντικειμενοστραφής προσομοιωτής που επιτρέπει:

- Να λαμβάνονται υπόψη διάφορα σενάρια διάδοσης και διασποράς σφαλμάτων (fault-propagation scenarios), όπως είναι για παράδειγμα η περίπτωση που τα αντίγραφα του αντικειμένου βρίσκονται στο ίδιο MT-domain.
- Την χρήση εναλλακτικών κατανομών εμφάνισης–επανάκτησης από λάθος (fault-repair distributions),
- Την εφαρμογή πιο ρεαλιστικών μοντέλων σφαλμάτων που εξαρτώνται από τον φόρτο εργασίας (load-dependent fault models), όπως αυτών που αναφέρονται στο [8].

ΚΕΦΑΛΑΙΟ 5ο

Λειτουργικότητα του Πρωτότυπου Προσομοιωτή

Σε αυτό το κεφάλαιο περιγράφεται η λειτουργικότητα του πρωτότυπου εργαλείου που υλοποιεί την προσέγγιση αξιολόγησης που θα παρουσιαστεί αργότερα.

5.1 Πολυνηματισμός

Σε πολλά κατανεμημένα συστήματα, είναι δυνατό να υπάρχουν πολλά νήματα ελέγχου (threads). Η ιδιότητα αυτή παρέχει ορισμένα σημαντικά πλεονεκτήματα, αλλά εισάγει και διάφορα προβλήματα.

Στο ανεπτυγμένο λογισμικό της προσομοίωσης υλοποιείται η ντετερμινιστική μη εκχωρίσιμη προσέγγιση ανάθεσης (non-preemptive deterministic scheduling) που περιγράφεται στο [18]. Επιβάλλεται από ένα λογικό νήμα ελέγχου για κάθε ένα πλεονασματικής επεξεργασίας αντικείμενο εξυπηρέτησης ή εξυπηρετούμενου κι εφαρμόζεται ντετερμινιστική ανάθεση των νημάτων και των μεθόδων επί των αντιγράφων κάθε αντικειμένου. Αυτό επιτρέπει την υποστήριξη ενός μεγάλου φάσματος μοντέλων ταυτοχρονισμού ([27]) όπως για παράδειγμα το νήμα-ανά-αίτηση

μοντέλο (thread-per-request model), το νήμα-ανά-σύνοδο μοντέλο (thread-per-session model), το νήμα-ανά-αντικείμενο μοντέλο (thread-per-object model) και το μοντέλο δεξαμενής νημάτων (thread pool model).

Στην τρέχουσα έκδοση του πρωτότυπου λογισμικού υλοποιείται το νήμα-ανά-αντικείμενο μοντέλο ταυτοχρονισμού, όπου ένα μοναδικό νήμα εκτελεί όλες τις μεθόδους ενός αντικειμένου. Στην συγκεκριμένη σχεδιαστική προσέγγιση, όμως, μπορεί να ενσωματωθεί κι ο παράγοντας της πολυνημάτωσης. Η ενσωμάτωση πολλαπλών νημάτων αποτελεί έναν από τους στόχους της επέκτασης του λογισμικού.

5.2 Εξισορρόπηση φόρτου

Θεωρείται ότι η διαχείριση συνέπειας κατάστασης (state consistency management) μεταξύ των ομάδων αντικειμένων στις οποίες ανατίθενται οι αιτήσεις εξυπηρέτησης, παραβιάζει την αρχή της διαφάνειας, εξ' αιτίας της εξάρτησης από την εκάστοτε εφαρμογή. Έτσι, περιοριζόμαστε σε στρατηγικές εξισορρόπησης φόρτου που δεν επιβάλουν κόστη διαχείρισης συνέπειας. Η απόφαση ανάθεσης αίτησης σε συγκεκριμένο αντικείμενο εξυπηρέτησης μπορεί να βασίζεται σε διαφορετικά επίπεδα διαμοιρασμού φόρτου: ανά αίτηση ή ανά σύνοδο.

- Μια *σύνοδος* (session) ορίζει την περίοδο κατά τη διάρκεια της οποίας ένας πελάτης είναι συνδεδεμένος με έναν εξυπηρετητή και στέλνει αιτήσεις σε κάποιο από τα αντικείμενα του εξυπηρετητή. Στον ανά σύνοδο διαμοιρασμό φόρτου, η διαχείριση επακόλουθων κλήσεων της ίδιας συνόδου γίνεται στο ίδιο αντικείμενο εξυπηρέτησης, το οποίο υποτίθεται ότι διαθέτει από μια κατάσταση για κάθε ανοιγμένη σύνοδο και οι άλλες σύνοδοι δεν εξαρτώνται, ούτε επηρεάζονται από αυτήν την κατάσταση. Αυτό το είδος εφαρμογών υπάρχει για παράδειγμα σε ένα κέντρο δικτύου επικοινωνίας που αναθέτει κυκλώματα για κάθε εισερχόμενη τηλεφωνική κλήση. Κρίσιμα δεδομένα συνόδου, όπως και στην ανάθεση της γραμμής, πρέπει να προστατεύονται από μια κατάλληλη πλεονασματικής επεξεργασίας πολιτική που θα εκτελεί την διαδοχή των λειτουργιών κλήσης-κλείσιμο συνδιαλλαγής των αιτημάτων. Αυτό το είδος ανά συνόδου ανάθεσης των αιτήσεων δεν

απαιτεί τον συγχρονισμό κατάστασης μεταξύ των διαθέσιμων αντικειμένων εξυπηρέτησης.

- Ο *ανά αίτηση* διαμοιρασμός φόρτου σε χωρίς καταστάσεις αντικείμενα εξυπηρέτησης έχει πρόσφατα γίνει μια συνηθισμένη επιλογή σχεδίασης σε 3- ή πολύ-επίπεδες αρχιτεκτονικές. Αυτές οι αρχιτεκτονικές ταιριάζουν στην λογική αποσύνθεση της δομής του υλικού και του λογισμικού των εφαρμογών σε εμφάνιση, λογική και δεδομένα. Οι εφαρμογές εξυπηρέτησης επεξεργάζονται τις αιτήσεις, αποθηκεύουν την κατάσταση (ή και τα αποτελέσματα), που προκύπτει μετά από την επεξεργασία, στο τέλος της ουράς καταγραφής (message log) και επιστρέφουν το αποτέλεσμα. Σε αντικείμενα εξυπηρέτησης χωρίς κατάσταση δεν υπάρχει ανάγκη για διαδικασίες τοποθέτησης σημείων καταγραφής κατάστασης, αντίθετα από τα αντικείμενα που διαθέτουν κατάσταση.

Οι υλοποιημένες, στο λογισμικό της εργασίας, στρατηγικές ανά σύνοδο ή ανά αίτησης διαμοιρασμού φόρτου είναι:

- τυχαία πιθανοτική (random probabilistic) με ίσες πιθανότητες για κάθε αντικείμενο εξυπηρέτησης (PROB)
- βασισμένη σε κατώφλι (threshold based - TB), όπου η ανάθεση αίτησης/ συνόδου σε κάποιο άλλο αντικείμενο εξυπηρέτησης ενεργοποιείται όταν ο αριθμός των αιτημάτων/ συνόδων στην ουρά ξεπεράσει ένα καθορισμένο πλήθος-κατώφλι.
- εκ περιτροπής ανάθεση (round robin - RR)
- ανάθεση στο αντικείμενο εξυπηρέτησης με τον μικρότερο αριθμό αιτημάτων/ ανοιγμένων συνόδων στην ουρά (EQT).

```
N ← NO_OF_SERVERS;
T ← THRESHOLD;
server ← choose with probability 1/N;
i ← server;
if (i.queue_length+1>T) {
 i←(i+1)%N;
 while (i.queue_length+1>T &&
i!=server)
 {
       i←(i+1)%NS;
 };
 server ← i;
}
```

(A)

```
N ← NO_OF_SERVERS;
server ← last used server;
k ← (server+1)%N;
i ← k;
ql ← 9999;
do {
if (i==k) {
       server ← i;
       ql ← server.queue_length;
}
else {
  if (server.primary.state==RECOVERING)
      && i.primary.state!=RECOVERING){
       server ← i;
       ql ← server.queue_length;}
  if (i.queue_length<ql
      && i.primary.state!=RECOVERING){
       server ← i;
       ql ← server.queue_length;}
}
i←(i+1)%N;

} while (i!=k);
```

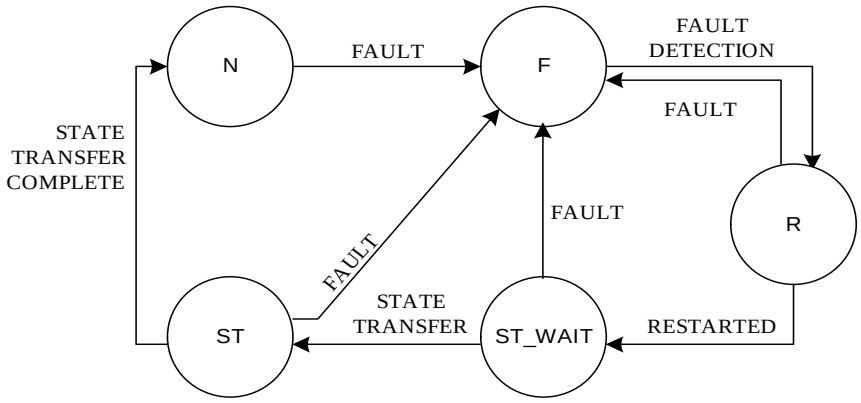
(B)

Σχήμα 4. Η TB (A) και η EQL (B) στρατηγικές ανάθεσης αιτημάτων.

5.3 Είδη Πλεονασματικής Επεξεργασίας

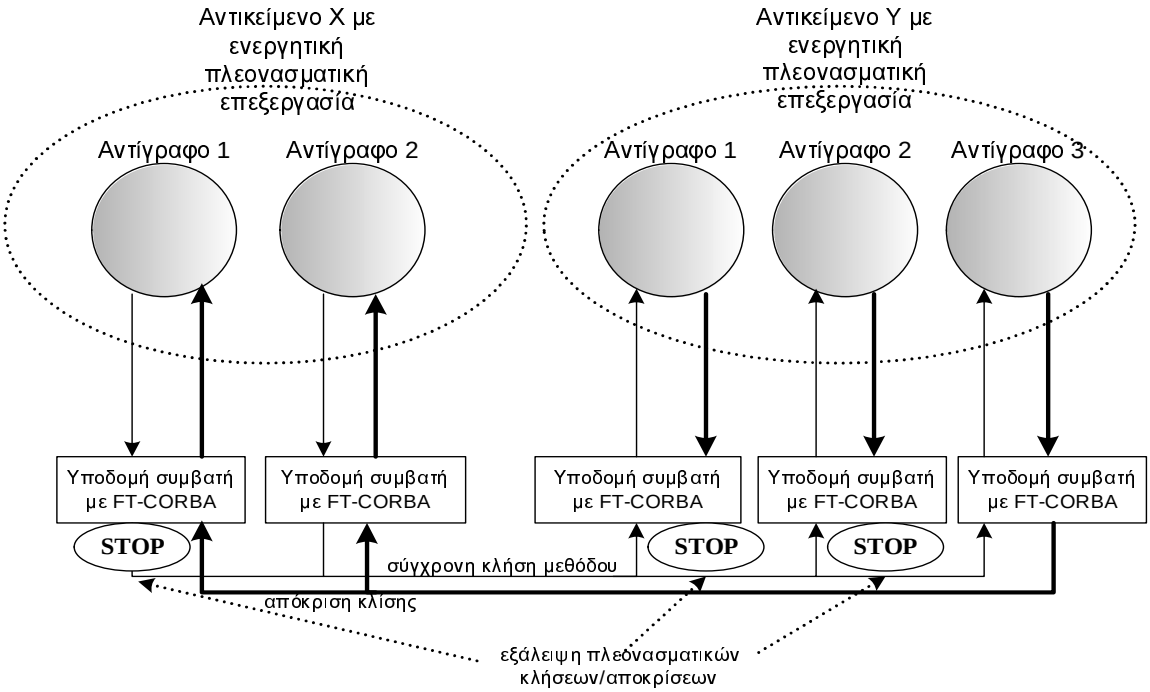
Σε αυτό το τμήμα, παρέχεται λεπτομερής παρουσίαση των παραμέτρων συμπεριφοράς των υλοποιημένων ειδών πλεονασματικής αντιγραφής.

Στην *ενεργητική πλεονασματική επεξεργασία* όλα τα μέλη της ομάδας αντικειμένων εκτελούν την κάθε αίτηση ανεξάρτητα, αλλά με την ίδια σειρά (κατάσταση N στο Σχήμα 5). Το κάθε μέλος διατηρεί ακριβώς την ίδια κατάσταση με τα άλλα μέλη, και στην περίπτωση του σφάλματος σε ένα μέλος (κατάσταση F στο Σχήμα 5) η προσομοιωμένη εφαρμογή συνεχίζει με τα αποτελέσματα να παρέχονται από τα άλλα μέλη, χωρίς να χρειαστεί να περιμένουμε την ανίχνευση του λάθους και την επανάκτησή του (κατάσταση R στο Σχήμα 5).



Σχήμα 5. Διάγραμμα μεταβάσεων των καταστάσεων ενός αντικειμένου ενεργής πλεονασματικής επεξεργασίας.

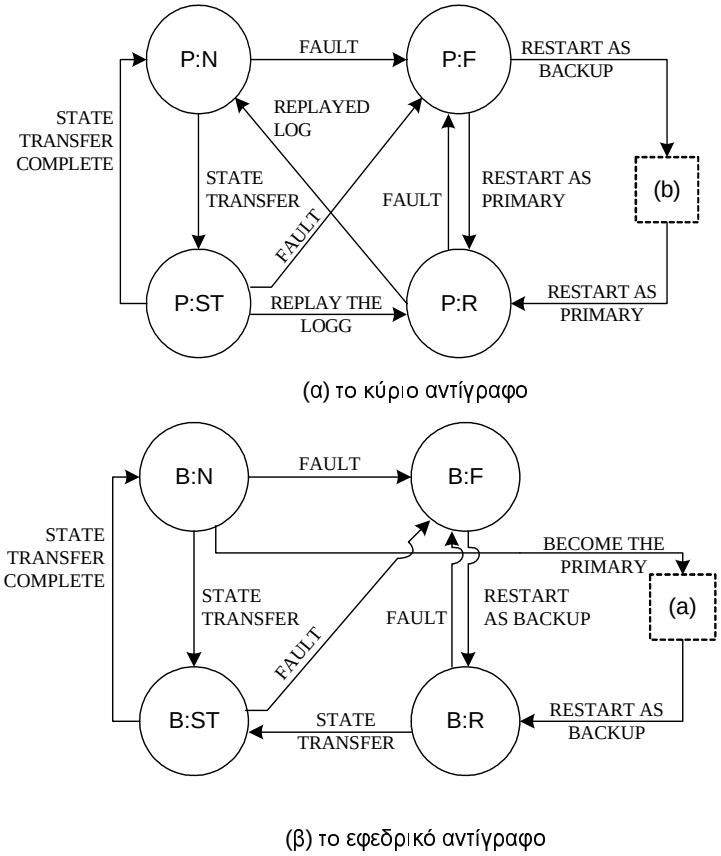
Ισχυρή συνέπεια ομάδας αντικειμένων στην ενεργητική πλεονασματική επεξεργασία σημαίνει ότι στο τέλος της λειτουργίας κάθε κλήσης μεθόδου, στην ομάδα αντικειμένων όλα τα μέλη της ομάδας έχουν την ίδια κατάσταση. Κάθε μέλος της ομάδας ανταποκρίνεται σε όλες τις εισερχόμενες αιτήσεις, γι' αυτό πιθανές διπλές αιτήσεις / απαντήσεις που μπορεί να προκληθούν, εντοπίζονται και καταστέλλονται, παραδίδοντας μια μόνο αίτηση / απάντηση στο αντικείμενο προορισμού.



Σχήμα 6. Σύγχρονη κλήση μεθόδου σε αντικείμενο ενεργής πλεονασματικής επεξεργασίας.

Η διαδικασία αντιγραφής κατάστασης (state transfer) μπορεί να καθυστερήσει μέχρι το αντίγραφο να αναρρώσει. Στο τέλος, λοιπόν, της επανάκτησης μια αντιγραφή (και καταγραφή) κατάστασης από ένα ενεργό (ελεύθερο από σφάλμα) μέλος της ομάδας αντικειμένων επιτυγχάνεται (κατάσταση ST στο Σχήμα 5). Τέτοιου είδους αντιγραφή κατάστασης απαιτεί λειτουργική απραξία (operational quiescence) και στα δυο μέλη της ομάδας αντικειμένων. Αυτό σημαίνει ότι η αντιγραφή της κατάστασης αναβάλλεται ενόσω όλα τα υπόλοιπα μέλη της ομάδας αντικειμένων βρίσκονται σε διαδικασία εξυπηρέτησης κάποιας αίτησης. Η αντιγραφή κατάστασης θα γίνει από το πρώτο μέλος που θα ολοκληρώσει την εξυπηρέτηση. Κατά την διάρκεια της επανάκτησης και της αντιγραφής κατάστασης, τα μέλη της ομάδας αντικειμένων είναι πιθανό να λαμβάνουν επιπλέον αιτήματα προς εξυπηρέτηση. Τα αιτήματα αυτά αποθηκεύονται τοπικά σε κάθε μέλος και σταδιακά αργότερα εξυπηρετούνται.

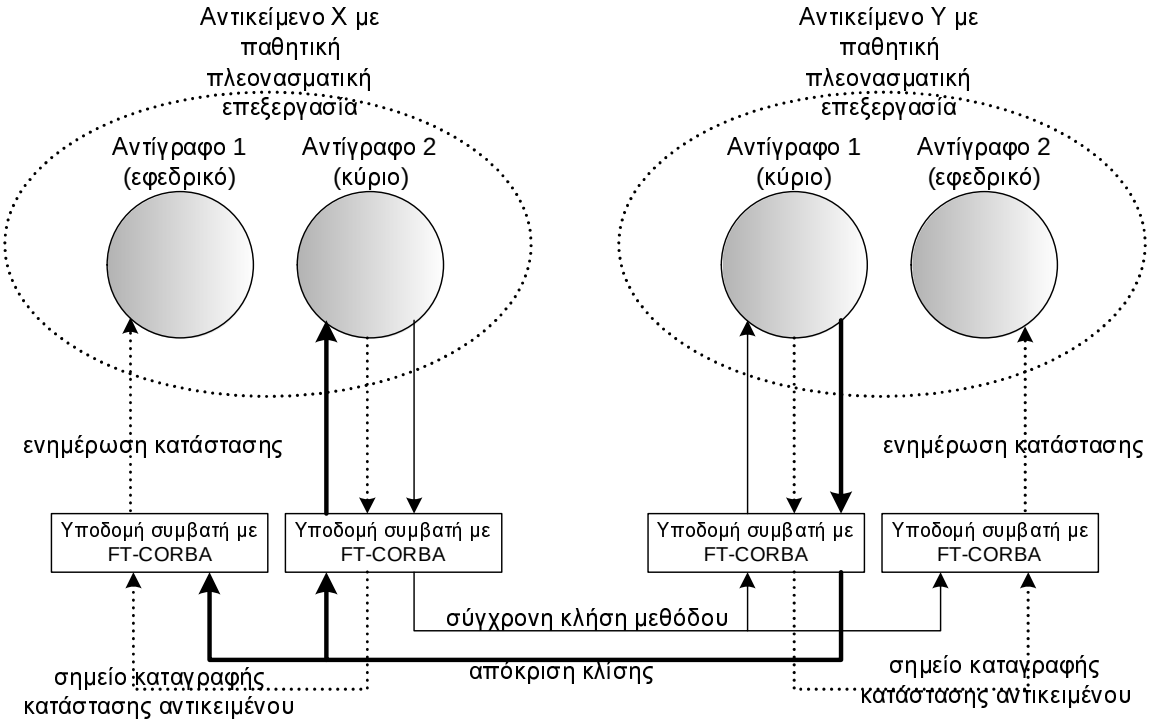
Στην *ψυχρή ή* στην *θερμή παθητική πλεονασματική επεξεργασία* κατά την διάρκεια της ελεύθερης από λάθη λειτουργίας μόνο ένα μέλος της ομάδας αντικειμένων, το κύριο αντίγραφο (primary, βλ. σχήμα 7a), εξυπηρετεί τις αιτήσεις που καταφθάνουν στην ομάδα. Η κατάσταση του κύριου αντιγράφου και η ακολουθία των μεθόδων καταγράφονται σε μια ουρά καταγραφής (message log), σύμφωνα με τις καθορισμένες τιμές που έχουν αποδοθεί στις παραμέτρους που αφορούν τα σημεία καταγραφής κατάστασης.



Σχήμα 7. Διάγραμμα μεταβάσεων των καταστάσεων ενός αντικειμένου θερμής παθητικής πλεονασματικής επεξεργασίας με ένα κύριο κι ένα εφεδρικό αντίγραφο.

Ισχυρή συνέπεια ομάδας αντικειμένων στην παθητική πλεονασματική επεξεργασία σημαίνει ότι στο τέλος κάθε σημείου καταγραφής ή αντιγραφής κατάστασης (μεταβάσεις P:ST -> P:N και B:ST -> B:N, στο Σχήμα 7), όλα τα μέλη της ομάδας είτε έχουν πρόσβαση στην ίδια κατάσταση είτε έχουν την ίδια κατάσταση. Στην παρουσία λαθών (κατάσταση P:F στο Σχήμα 7), ένα εφεδρικό αντίγραφο προωθείται να γίνει το νέο κύριο (μετάβαση B:N -> P:R στο Σχήμα 7). Ως κατάσταση του νέου κύριου αντιγράφου φορτώνεται η τελευταία σωσμένη κατάσταση του προηγούμενου κύριου αντιγράφου. Ύστερα, το νέο κύριο αντίγραφο εξυπηρετεί ξανά τις αιτήσεις, με την ίδια ακριβώς σειρά, που είχαν αποθηκευτεί στην ουρά καταγραφής μετά το τελευταίο σημείο καταγραφής κατάστασης. Έτσι, όταν θα σταματήσει αυτή η εκτέλεση των μεθόδων, το κύριο αντίγραφο θα βρίσκεται ακριβώς στην ίδια κατάσταση στην οποία βρισκόταν το προηγούμενο κύριο αντίγραφο όταν συνέβη το σφάλμα. Αυτό σημαίνει ότι ένας πελάτης μπορεί να ξαναστείλει μια αίτηση σε έναν εξυπηρετητή και να λάβει κι απάντηση σε αυτό το αίτημα ή ότι ένας εξυπηρετητής μπορεί να στείλει διπλή μια

απάντηση. Όπως όμως αναφέρθηκε και στην παράγραφο 4.2, διπλές, ή και πολλαπλές, κλήσεις που μπορεί να υπάρξουν εντοπίζονται και καταστέλλονται, κι έτσι δεν υπάρχει κίνδυνος μια μέθοδος να εκτελεστεί περισσότερο από μια φορά.



Σχήμα 8. Σύγχρονη κλήση μεθόδου σε αντικείμενο παθητικής πλεονασματικής επεξεργασίας που κάνει χρήση σημείων καταγραφής κατάστασης.

Στην ψυχρή παθητική πλεονασματική επεξεργασία, τα εφεδρικά αντίγραφα του αντικειμένου δεν είναι καν ενεργοποιημένα. Όταν στο τρέχων κύριο αντίγραφο συμβαίνει κάποιο λάθος, ένα άλλο επιλέγεται και ενεργοποιείται. Στην θερμή παθητική πλεονασματική επεξεργασία τα εφεδρικά αντίγραφα είναι ήδη ενεργοποιημένα και οι καταστάσεις τους συγχρονίζονται συνεχώς με αυτήν του κύριου αντιγράφου σύμφωνα με την προκαθορισμένη συχνότητα.

Η κατανάλωση πόρων για την καταγραφή και την αντιγραφή κατάστασης εξαρτάται από το μέγεθος της κατάστασης του αντικειμένου κι από την ταχύτητα επεξεργασίας του πιο αργού συμμετέχοντος στη διαδικασία αντιγράφου της ομάδας αντικειμένων. Μια καταγραφή ή αντιγραφή κατάστασης επιτρέπεται να αρχίσει μόνο όταν δεν παραβιάζεται η αρχή της ισχυρής συνέπειας ομάδας αντικειμένων. Έτσι, αναβάλλεται όταν το κύριο αντίγραφο βρίσκεται εν μέσω εξυπηρέτησης ή παραμένει αδρανές

περιμένοντας να λάβει κάποια απάντηση. Κατά την διάρκεια της διαδικασίας της καταγραφής ή της αντιγραφής κατάστασης, είναι πιθανό νέες αιτήσεις προς εξυπηρέτηση να καταφθάνουν, όμως η εξυπηρέτησή τους δεν μπορεί να αρχίσει μέχρι να τελειώσει η διαδικασία.

5.4 Πολιτικές Αναμονής Απόκρισης – Επανάκλησης

Το ανεπτυγμένο πρωτότυπο λογισμικό προσομοίωσης υποστηρίζει επαναλαμβανόμενες κλήσεις μεθόδου (request re-invocation) ως μέσο αναγνώρισης λαθών στον παραλήπτη ή λαθών δικτύου. Και στα δυο είδη λαθών ο αιτούμενος την εξυπηρέτηση (πελάτης) δεν μπορεί να ανιχνεύσει το πρόβλημα και δεν λαμβάνει απάντηση.

Μια τυπική λύση, σε τέτοιες περιπτώσεις, είναι η χρήση, στην εφαρμογή του εξυπηρετούμενου, διαστημάτων επανάληψης (timeouts). Εναλλακτικά, το πρότυπο OMG FT-CORBA προτείνει την χρήση μιας κατάλληλης πολιτικής «παλμών» (heart-beat policy): ο πελάτης χρησιμοποιεί την ίδια σύνοδο για να ενεργοποιεί ξανά την αίτηση τόσο συχνά όσο προσδιορίζεται από τις παραμέτρους διαστήματος παλμού. Αν η αντίστοιχη απάντηση δεν φτάσει εντός κάποιου προκαθορισμένου διαστήματος παλμού, νέα σύνοδος χρησιμοποιείται για το χαμένο αίτημα.

Στο πρωτότυπο λογισμικό προσομοίωσης της εργασίας και οι δυο μηχανισμοί μοντελοποιούνται με την χρήση διαστημάτων αναμονής απόκρισης – επανάκλησης (re-quest-retry timeouts). Στο λογισμικό λαμβάνονται επίσης υπόψη και τα συνοδευτικά επιπλέον κόστη επαναποστολής των (πιθανώς χαμένων) αιτημάτων.

5.5 Ανίχνευση Λαθών

Το μοντελοποιημένο περιβάλλον ανίχνευσης λαθών υποθέτει την ύπαρξη μιας διαφανούς και με ανοχή σε λάθη υπηρεσίας παρακολούθησης (transparent and fault tol-erant *monitoring service*). Γενικά, ο ανιχνευτής σφάλματος που παρακολουθεί ένα αντικείμενο μιας κάποιας εφαρμογής συνήθως βρίσκεται, για καλύτερη αποτελεσματικότητα, στον ίδιο υπολογιστικό σταθμό (host) με το αντικείμενο.

Επιπλέον, ένας γενικός ανιχνευτής (global detector) που χρησιμοποιεί πολιτικές πλεονασματικής επεξεργασίας, ώστε να είναι ανεκτικός σε λάθη, παρακολουθεί τους τοπικούς ανιχνευτές λαθών.

Έχει βρεθεί ([7]) ότι η διαδικασία παρακολούθησης προκαλεί περίπου 5% προσαύξηση στην χρησιμοποίηση του επεξεργαστή (ενός Pentium-II 200+Hz υπολογιστή), για περίπου 500 msec. Αυτό το κόστος το λαμβάνει υπόψη το ανεπτυγμένο πρωτότυπο λογισμικό προσομοίωσης της εργασίας, παρόλο που δεν χρησιμοποιεί κάποιον συγκεκριμένο μηχανισμό παρακολούθησης για ανίχνευση λαθών από αυτούς που περιγράφονται στο πρότυπο OMG FT-CORBA. Οι δυο μηχανισμοί, παρακολούθησης για την ανίχνευση λαθών, που περιγράφονται στο πρότυπο αυτό διαφέρουν μόνο ως προς την κατεύθυνση προς την οποία ρέει μέσα στο σύστημα η πληροφορία σχετικά με τα λάθη.

Η υλοποίηση στο λογισμικό προσομοίωσης της εργασίας *προβλέπει τον περιοδικό έλεγχο κάθε αντικειμένου* (της ομάδας αντικειμένων και των μελών αυτής), σύμφωνα με καθορισμένο χρονικό διάστημα το οποίο:

- στην περίπτωση της pull-based παρακολούθησης αναπαριστά το άθροισμα του διαστήματος παρακολούθησης και του χρόνου που χρειάζεται για να φτάσει η απάντηση από το αντικείμενο, η οποία προσδιορίζει αν στο αντικείμενο αυτό έχει συμβεί σφάλμα ή όχι.
- στην περίπτωση της push-based παρακολούθησης αναπαριστά τον χρόνο που επιτρέπεται στο αντικείμενο το οποίο παρακολουθείται να δηλώσει αν του έχει συμβεί λάθος ή όχι.

ΚΕΦΑΛΑΙΟ 6ο

Η Προσέγγιση Αξιολόγησης (Evaluation Approach)

6.1 Εισαγωγή

Σε αυτό το τμήμα της εργασίας παρουσιάζεται η προσέγγιση για την αξιολόγηση της απόδοσης των πιο σημαντικών παραγόντων αντιστάθμισης μεταξύ της απόδοσης και της αποτελεσματικότητας της ανοχής λαθών (fault-tolerance performance and effectiveness tradeoffs) που εμφανίζονται σε *σύνθετα σχήματα πλεονασματικής επεξεργασίας* (composite replication schemes).

Όπως έχει ήδη αναφερθεί, μια δομή παροχής ανοχής σε σφάλματα (fault tolerance setting) που αποτελείται από αντικείμενα (objects) τα οποία πιθανώς ακολουθούν διαφορετικές πολιτικές πλεονασματικής επεξεργασίας (replication policies) ονομάζεται *σύνθετο σχήμα πλεονασματικής επεξεργασίας* (composite replication scheme). Το είδος πλεονασματικής αντιγραφής (replication style) που ανατίθεται σε κάθε αντικείμενο μπορεί να είναι *ενεργό* (active replication), *θερμό παθητικό* (warm passive replication) ή *ψυχρό παθητικό* (cold passive replication). Το σύνολο των ιδιοτήτων που καθορίζουν την συμπεριφορά του είδους της πλεονασματικής επεξεργασίας (όπως για παράδειγμα το πλήθος των αντιγράφων, ο καθορισμός των σημείων καταγραφής κατάστασης, τα διαστήματα αντιγραφής κατάστασης, τα διαστήματα αναμονής απόκρισης-επανάκλησης, κτλ.) πρέπει να παίρνουν τιμές κατάλληλες, σε σχέση με την εφαρμοζόμενη δομή ανίχνευσης των σφαλμάτων.

Η προτεινόμενη προσέγγιση μπορεί να εφαρμοστεί ανεξάρτητα από το επίπεδο αφαίρεσης του μοντέλου που προσομοιώνεται. Το επίπεδο αφαίρεσης αποφασίζεται με βάση:

- την ανάγκη να ελαχιστοποιηθούν το σύνολο των παραμέτρων του μοντέλου και το σύνολο των διαφορετικών τύπων γεγονότων που πρέπει να προσομοιωθούν, ενώ την ίδια στιγμή θέλουμε
- την παροχή αξιόπιστων αποτελεσμάτων που είτε παρέχουν διορατικότητα για τους παρόντες παράγοντες αντιστάθμισης είτε αποτελούν ακριβείς προβλέψεις για την απόδοση (performance) και την αποτελεσματικότητα (effectiveness) της ανοχής σε σφάλματα που θα παρείχε το πλεονασματικής επεξεργασίας σύστημα που προσομοιώνεται.

Η αξιοπιστία των παραγόμενων αποτελεσμάτων εξασφαλίζεται από την εφαρμοζόμενη ανάλυση των αποτελεσμάτων της προσομοίωσης (βλ. παράγραφο 6.2). Η τελική ακρίβεια που λαμβάνουμε, όπως και σε κάθε μελέτη βασισμένη σε προσομοίωση, εξαρτάται από το επίπεδο αφαίρεσης του μοντέλου κι από την διαθεσιμότητα των κατάλληλων πληροφοριών σχετικά με την κατανάλωση πόρων και την ύπαρξη λαθών: τα παραγόμενα αποτελέσματα πρέπει να αξιολογούνται σύμφωνα με την έκδοση του λογισμικού της προσομοίωσης του συστήματος που αναπτύχθηκε και χρησιμοποιήθηκε.

Στην περίπτωση της παρούσης εργασίας, το επίπεδο της αφαίρεσης του μοντέλου μπορεί να γίνει αντιληπτό από το σύνολο των παραμέτρων που χρησιμοποιούνται (βλ. κεφάλαιο 7). Το ενδιαφέρον της εργασίας επικεντρώνεται κυρίως στην προσέγγιση της εκτίμησης της απόδοσης, ενώ η αξιολόγηση του μοντέλου που χρησιμοποιήθηκε δεν απασχολεί ιδιαίτερα. Η εργασία τονίζει την δυνατότητα της συγκεκριμένης προσέγγισης να παρέχει διορατικότητα για πολύ σημαντικά ζητήματα, όπως οι παράγοντες αντιστάθμισης μεταξύ της απόδοσης και της αποτελεσματικότητας της ανοχής σε λάθη και την δυνατότητα της, δυνητικά, να υποστηρίζει συστηματικές μεθόδους σχεδιασμού ποιότητας εξυπηρέτησης (systematic QoS design methods), για τον καθορισμό είτε

- εγγυήσεων για τους χρόνους απόκρισης είτε
- της βέλτιστης παραμετροποίησης συστήματος με ανοχή σε σφάλματα (καθορισμός των σημείων καταγραφής κατάστασης, των διαστημάτων αντιγραφής κατάστασης, αναμονής απόκρισης-επανάκλησης, κτλ.), με στόχο την επίτευξη συγκεκριμένων επιπέδων ποιότητας εξυπηρέτησης.

Ο πρωτότυπος προσομοιωτής που αναπτύχθηκε επιτρέπει την μοντελοποίηση των αλληλεπιδράσεων μεταξύ των αντικειμένων όσον αφορά:

- την ταυτόχρονη κατοχή πόρων (simultaneous resource possession) που προκαλείται από τις σύγχρονες (synchronous) και συχνά εμφωλιασμένες κλήσεις μεθόδων
- τον ανταγωνισμό κατοχής των υλικών πόρων (hardware resource contention), ως αποτέλεσμα της τοποθέτησης των μελών των διαφόρων ομάδων αντικειμένων στα διάφορα υπολογιστικά κέντρα
- τον επιπλέον φόρτο και τα επιπλέον κόστη (requests' blocking costs) που προκαλούνται από την αναβολή εξυπηρέτησης αιτήσεων εξαιτίας των διαφόρων (αν υπάρχουν) δραστηριοτήτων καταγραφής ή /και αντιγραφής καταστάσεων
- τον επιπλέον φόρτο που προκαλείται από την επανεκκίνηση (replica restart) κάποιου μέλους μιας ομάδας αντικειμένων ή από την επανεπεξεργασία των αποθηκευμένων στην ουρά καταγραφής αιτήσεων (log replay)
- το επιπλέον κόστος που προκαλείται από επανάκληση (re-invocation) αιτήσεων οι οποίες πιθανώς χάθηκαν είτε λόγω σφάλματος δικτύου είτε λόγω λάθους στον παραλήπτη της κλήσης
- το επιπλέον κόστος που προκαλείται από τον περιοδικό έλεγχο παρακολούθησης (monitoring) αν σε ένα αντικείμενο έχει συμβεί λάθος ή όχι.

6.2 Απόδοση και Αποτελεσματικότητα της Ανοχής σε Λάθη

Για τον καθορισμό των μέτρων της αξιολόγησης και της αποτελεσματικότητας της ανοχής σε λάθη, γίνεται ξεκάθαρος διαχωρισμός μεταξύ

- των αιτημάτων που δεν έχουν επηρεαστεί από τα λάθη που συμβαίνουν (fault unaffected requests) και
- των αιτημάτων που έχουν επηρεαστεί από τα λάθη που συμβαίνουν (fault affected requests)

Οι μη επηρεασμένες από τα λάθη αιτήσεις αποτελούν την συντριπτική πλειοψηφία των διακινούμενων αιτήσεων, αλλά σε ένα αξιόπιστο σύστημα οι εγγυήσεις για την ποιότητα της εξυπηρέτησης πρέπει να συμπεριλαμβάνουν και τους χρόνους απόκρισης των επηρεασμένων από τα λάθη αιτήσεων.

Ορισμός: Μια αίτηση εξυπηρέτησης (δεν) είναι επηρεασμένη από τα λάθη αν (δεν) έχει καθυστερήσει η εξυπηρέτησή της, ή / και αν η αποστολή της στον εξυπηρετητή (δεν) έχει καθυστερήσει, εξαιτίας τουλάχιστον ενός ανιχνευμένου λάθους.

Λαμβάνοντας υπόψη την αξιόπιστη πλήρως διατεταγμένη παράδοση των αιτήσεων (βλ. παράγραφο 4.1) σε όλα τα μέλη της ίδιας ομάδας αντικειμένων, δίνονται οι παρακάτω ορισμοί:

Ορισμός: Οποιαδήποτε σύγχρονη αίτηση προς ένα αντικείμενο ενεργής πλεονασματικής επεξεργασίας θεωρείται πως είναι επηρεασμένη από λάθη αν η εξυπηρέτησή της καθυστερείται :

- από αναμονή (blocking) μέχρι να ικανοποιηθεί λειτουργική απραξία (operational quiescence) και/ ή
- από αντιγραφή κατάστασης (state transfer) κατά την διάρκεια της επανάκτησης (recovery) μέλους ομάδας αντικειμένων ή
- επειδή έχει τοποθετηθεί σε ουρά πίσω από κάποια άλλη αίτηση, η οποία όμως έχει επηρεαστεί από λάθος.

Σε ένα αντικείμενο ενεργής πλεονασματικής επεξεργασίας, σε κάθε μέλος της ίδιας ομάδας αντικειμένων μοιράζεται από ένα αντίγραφο της κάθε αίτησης που καταφθάνει. Αν, σε ένα τέτοιο αντικείμενο, τουλάχιστον ένα από τα μέλη της ομάδας έχει εξυπηρετήσει το αντίστοιχο αντίγραφο της αίτησης και έχει αποστείλει και απάντηση προτού συμβεί κάποιο λάθος, τότε η αίτηση αυτή δεν είναι επηρεασμένη από λάθος (αφού δεν καθυστέρησε εξαιτίας του λάθους).

Ορισμός: Οποιαδήποτε αίτηση προς ένα παθητικής πλεονασματικής επεξεργασίας αντικείμενο θεωρείται πως είναι επηρεασμένο από λάθη, αν

- έχει καθυστερήσει επειδή χάθηκε από ή βρέθηκε σε ουρά κύριου αντιγράφου (primary) ομάδας αντικειμένων που ήρθε σε κατάσταση λάθους (failed) ή επανάκτησης (recovering) ή

- έχει τοποθετηθεί σε ουρά πίσω από κάποια άλλη αίτηση, η οποία όμως έχει επηρεαστεί από λάθος.

Ως συνέπεια των προηγούμενων ορισμών, μια αίτηση εξυπηρέτησης επηρεάζεται από τα λάθη που συμβαίνουν, αν τουλάχιστον μια από τις κλήσεις της επηρεαστεί ή αν αποθηκευτεί σε ουρά πίσω από άλλες αιτήσεις που έχουν επηρεαστεί από λάθη. Όλες οι υπόλοιπες αιτήσεις θεωρούνται μη επηρεασμένες από λάθη. Τα μόνα από τα οποία επηρεάζονται αυτές οι αιτήσεις, είναι τα επιπλέον κόστη κι ο επιπλέον φόρτος που επιβάλλονται από το είδος της πλεονασματικής επεξεργασίας που χρησιμοποιείται.

Τα μέτρα που χρησιμοποιούνται για τις μη επηρεασμένες από λάθη αιτήσεις, ποσοτικοποιούν την *απόδοση* (performance) του υπό προσομοίωση είδους πλεονασματικής επεξεργασίας. Από την άλλη όμως, τα μέτρα που χρησιμοποιούνται για τις επηρεασμένες από λάθη αιτήσεις, μπορούν να χρησιμοποιηθούν:

- είτε για να παρέχουν εγγυήσεις ποιότητας εξυπηρέτησης όσον αφορά τους χρόνους απόκρισης
- είτε για να ποσοτικοποιηθεί η *αποτελεσματικότητα* της ανοχής σε λάθη (fault tolerance effectiveness) του εφαρμοζόμενου είδους πλεονασματικής επεξεργασίας.

Σε υβριδικές προσομοιώσεις, αξιοπιστίας και κυκλοφορίας συστήματος, είναι επίσης πιθανό να παράγονται εκτιμήσεις τυπικής αξιοπιστίας και διαθεσιμότητας συστήματος εξυπηρέτησης. Ωστόσο, μετρήσεις βασισμένες στην κυκλοφορία των αιτήσεων και που ορίζονται ξεχωριστά για τις αιτήσεις που είναι επηρεασμένες από λάθη και ξεχωριστά για τις μη επηρεασμένες από λάθη αιτήσεις, είναι καλύτερες στο να παρέχουν την ουσία των πιο σημαντικών παραγόντων αντιστάθμισης μεταξύ της απόδοσης και της αποτελεσματικότητας της ανοχής λαθών.

Η προσέγγιση αξιολόγησης, που παρουσιάζεται, μπορεί να χρησιμοποιηθεί ως μέσο σχεδίασης συστήματος με τους ακόλουθους τρόπους:

- Για να καθοριστούν οι ελάχιστοι χρόνοι απόκρισης των επηρεασμένων από σφάλματα αιτημάτων, δηλαδή την *βέλτιστη αποτελεσματικότητα* (optimum effectiveness), που ένα σύνθετο σχήμα πλεονασματικής επεξεργασίας μπορεί να αποδώσει, για οποιονδήποτε πιθανό συνδυασμό τιμών ανατίθεται στο σύνολο των ιδιοτήτων που καθορίζουν την συμπεριφορά του σχήματος.

Σε συστήματα τα οποία χρησιμοποιούν σύνθετα σχήματα πλεονασματικής επεξεργασίας *η βέλτιστη αποτελεσματικότητά τους είναι το μοναδικό μέτρο που μπορεί να κάνει εφικτή την σύγκριση μεταξύ τους*. Έτσι, προτιμάται εκείνο το σύνθετο σχήμα πλεονασματικής επεξεργασίας, το οποίο ικανοποιεί το επιθυμητό επίπεδο ποιότητας εξυπηρέτησης με το μικρότερο κόστος, δηλαδή εκείνο το σύνθετο σχήμα πλεονασματικής επεξεργασίας το οποίο επιδεικνύει ανοχή σε λάθη με την βέλτιστη απόδοση (best fault tolerance performance).

- Για τον καθορισμό των κατάλληλων τιμών, που πρέπει να αποδοθούν στις παραμέτρους που υποδεικνύουν την συμπεριφορά, ώστε να επιτευχθεί κάποιο συγκεκριμένο επιθυμητό επίπεδο ποιότητας εξυπηρέτησης (όχι απαραίτητα εκείνο που δίνει την βέλτιστη αποτελεσματικότητα). Αυτό μπορεί να γίνει με την κατάλληλή ανάλυση των παραγόντων αντιστάθμισης, όπου με κάθε πιθανή αλλαγή των παραμέτρων σε κάποιο πλεονασματικό είδος επεξεργασίας (ή και με αλλαγή του ίδιου του είδους), *ανταλλάσσουμε τις δυνητικές βελτιώσεις στους χρόνους απόκρισης των επηρεασμένων από λάθη αιτήσεων με τα επιπλέον κόστη που επιβάλλονται στην εξυπηρέτηση των μη επηρεασμένων από λάθη αιτημάτων*. Για σύνθετα σχήματα πλεονασματικής επεξεργασίας, τέτοιου είδους ανάλυση συγκλίνει στον συνδυασμό που ικανοποιεί το σύνολο των σχεδιαστικών στόχων στο κατώτερο δυνατό κόστος.

Η ίδια προσέγγιση έχει ήδη χρησιμοποιηθεί στο [11], για να θεμελιώσει μια ολοκληρωμένη μέθοδο σχεδίασης ποιότητας εξυπηρέτησης (QoS design method). Η προτεινόμενη μέθοδος σχεδίασης στοχεύει στην επιλογή των κατάλληλων σημείων καταγραφής κατάστασης και των κατάλληλων διαστημάτων αντιγραφής κατάστασης σε αντικείμενα παθητικής πλεονασματικής επεξεργασίας. Οι βέλτιστες συνθήκες αποτελεσματικότητας, που καλούνται οριακές ρυθμίσεις αποτελεσματικότητας (tightest effective intervals), για τα υποψήφια σχήματα καθορίζονται από την στατιστική βελτιστοποίηση της προσομοίωσης [12], στα πλαίσια ενός κατάλληλα επιλεγμένου ενιαίου πειραματικού σχεδιασμού [29]. Δηλαδή, η προτεινόμενη στο [11] διαδικασία απόφασης, βασισμένη στους παράγοντες αντιστάθμισης, επιτρέπει την επιλογή των με τα πιο χαμηλά κόστη σημείων καταγραφής κατάστασης και των κατάλληλων

διαστημάτων αντιγραφής κατάστασης, με στόχο πάντα την εκπλήρωση του συνόλου των σχεδιαστικών στόχων.

Στο επόμενο Κεφάλαιο, η προσέγγιση αξιολόγησης που παρουσιάζει η παρούσα εργασία και το πρωτότυπο λογισμικό που αναπτύχθηκε χρησιμοποιούνται και παρέχουν διαίσθηση και για έναν άλλο σημαντικό παράγοντα αντιστάθμισης μεταξύ της απόδοσης και της αποτελεσματικότητας: στο γεγονός ότι η χρήση υπερβολικά σύντομων διαστημάτων αναμονής απόκρισης-επανάκλησης, από τα αντικείμενα που αποτελούν το σύστημα, μπορεί να προκαλέσει υψηλά επιπλέον κόστη και τελικά να μην βελτιώσει την αποτελεσματικότητα της ανοχής σε λάθη.

6.3 Η Ανάλυση των Αποτελεσμάτων της Προσομοίωσης

Ιδιαίτερη έμφαση έχει δοθεί στην αξιοπιστία των παραγόμενων μέσων όρων, με την χρήση κατάλληλων διαδικασιών ανάλυσης των αποτελεσμάτων.

Στην περίπτωση αντικειμένων παθητικής πλεονασματικής επεξεργασίας εφαρμόζεται προσομοίωση μιας μόνο εκτέλεσης (single-run procedure) ([20]), η οποία εκμεταλλεύεται την ποσοτική αναπαράσταση των απαιτούμενων εκτιμήσεων στην κατάσταση σταθερότητας (steady-state). Η κατάσταση σταθερότητας επιτυγχάνεται αφού η προσομοίωση (ή ακόμη και η ίδια η λειτουργία) ενός συστήματος έχει διαρκέσει αρκετό χρόνο ώστε το σύστημα να βρίσκεται πλέον σε μια μέση κατάσταση ισορροπίας, σε μια κατάσταση, δηλαδή, η οποία δεν επηρεάζεται άμεσα ή έμμεσα από τις αρχικές συνθήκες, αλλά αντίθετα έχει πια αποκτήσει σταθερή δυναμική. Η απόδοση σε αυτήν την κατάσταση σταθερότητας είναι που συνήθως ενδιαφέρει. Οι εκτιμήσεις σταθερής κατάστασης βασίζονται στις μετρήσεις που γίνονται μεταξύ δυο διαδοχικών εισόδων του συστήματος σε επιλεγμένο σύνολο καταστάσεων, έστω Α. Ένα τέτοιο σύνολο-στόχος καταστάσεων για τους μέσους χρόνους απόκρισης σε ένα αντικείμενο εξυπηρέτησης, περιλαμβάνει οποιαδήποτε κατάσταση κατά την οποία στο κύριο αντίγραφο συμβαίνει λάθος και το πλήθος των αποθηκευμένων στην ουρά αιτημάτων είναι μηδέν (0). Οι Α-κύκλοι δεν είναι μεταξύ τους ανεξάρτητοι κι ομοιόμορφα κατανεμημένοι (independent and identically distributed) και για αυτόν τον λόγο γίνεται

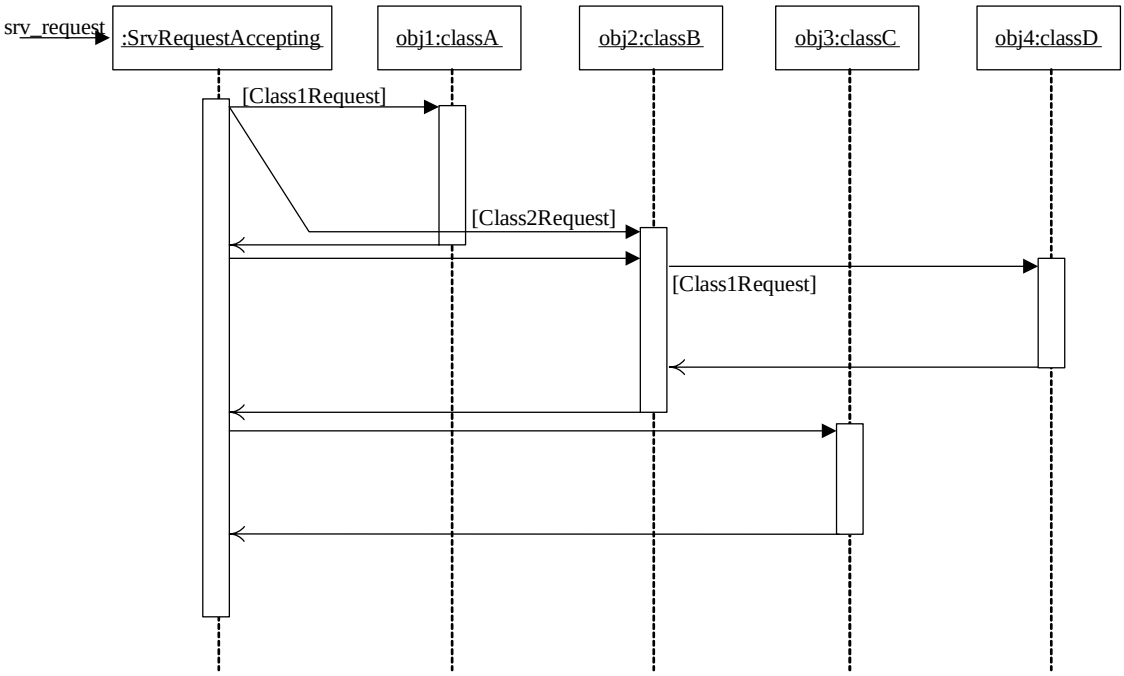
χρήση μαζικής εκτίμησης (batch means estimation). Οι ποσότητες που βασίζονται σε διαδοχικούς Α-κύκλους ομαδοποιούνται σε μη επικαλυπτόμενες παρτίδες και η μεταχείριση των μέσων τους γίνεται σαν να ήταν ανεξάρτητες κι ομοιόμορφες παρατηρήσεις. Η εγκυρότητα αυτής της προσέγγισης αυξάνει όσο αυξάνει και το μέγεθος της κάθε παρτίδας. Το πλήθος των Α-κύκλων καθώς και το μέγεθος της παρτίδας καθορίζονται δυναμικά, από την ακολουθιακή διαδικασία ελέγχου των Law και Carson (Law and Carson sequential control procedure) ([13]), στη βάση πάντα της καθορισμένης σχετικής ακρίβειας που πρέπει να επιτευχθεί.

Σε περιπτώσεις διαμόρφωσης συστήματος (π.χ. αντικείμενα εξυπηρέτησης που χρησιμοποιούν θερμή πλεονασματική επεξεργασία) ή καταμερισμού φόρτου εργασίας όπου δεν είναι εύκολο να ορισθεί ένα τέτοιο σύνολο καταστάσεων, στο οποίο το σύστημα να εισέρχεται σχετικά συχνά, γίνεται χρήση της πολύ γνωστής προσέγγισης των ανεξάρτητων πλεονασματικών αντιγράφων (independent replications approach).

ΚΕΦΑΛΑΙΟ 7ο

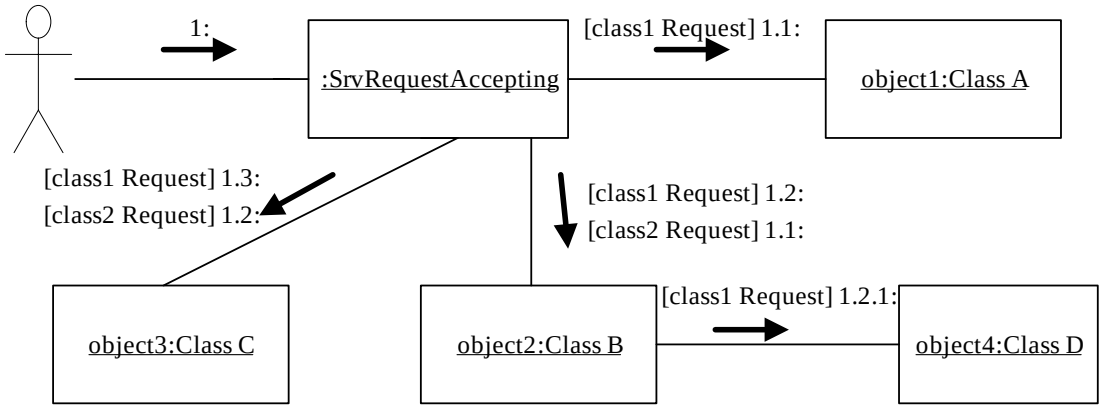
Μελέτη Περίπτωσης (Case System Study)

Σε αυτό το Κεφάλαιο της εργασίας παρουσιάζονται αποτελέσματα που σταδιακά «ξετυλίγουν» και με πολύ παραστατικό τρόπο περιγράφουν τους παράγοντες αντιστάθμισης που προκύπτουν στην προσπάθεια χρήσης πιο αποτελεσματικών (σύντομων) διαστημάτων αναμονής απόκρισης-επανάκλησης ώστε να αντιμετωπίζονται λάθη δικτύου ή άλλα λάθη (όπως αυτά που περιγράφονται στην παράγραφο 4.2). Σε ένα σύνθετο σχήμα πλεονασματικής επεξεργασίας υπερβολικά συχνά διαστήματα αναμονής απόκρισης-επανάκλησης προκαλούν υψηλά επιπλέον κόστη και δεν βελτιώνουν την αποτελεσματικότητα της ανοχής σε λάθη.



Σχήμα 9. Η ακολουθία των μηνυμάτων για τα αντικείμενα του συστήματος της μελέτης περίπτωσης.

Το μοντέλο συστήματος που χρησιμοποιήθηκε επιτρέπει την δοκιμή ενός φάσματος διαφορετικών συνδυασμών λαθών και πολιτικών αναμονής απόκρισης-επανάκλησης. Το μοντέλο συστήματος που χρησιμοποιήθηκε περιλαμβάνει ένα ενεργής πλεονασματικής επεξεργασίας αντικείμενο. Επίσης, αποτελείται από τέσσερα (4) αντικείμενα που δεν έχουν καταστάσεις (στιγμιότυπα της κλάσης SrvRequestAccepting) κι από τέσσερα (4) αντικείμενα που αλλάζουν καταστάσεις (obj1, obj2, obj3, obj4). Τα αντικείμενα αυτά αλληλεπιδρούν μεταξύ τους όπως φαίνεται στο Σχήμα 9 και στο Σχήμα 10. Χρησιμοποιούνται δυο (2) είδη αιτημάτων: αιτήσεις της κλάσης class1 και αιτήσεις της κλάσης class2. Κατά την άφιξή τους τα αιτήματα ανατίθενται στα διαθέσιμα αντικείμενα σύμφωνα με την εκ περιτροπής (Round Robin, RR) πολιτική.



Σχήμα 10. Το διάγραμμα συνεργασίας των αντικειμένων.

7.1 Η Δομή του Μοντέλου Συστήματος

Τα τέσσερα χωρίς κατάσταση αντικείμενα εξυπηρέτησης (obj0, obj5, obj6, obj7), καθώς και τα αντικείμενα obj2, obj3 και obj4 χρησιμοποιούν θερμή παθητική πλεονασματική επεξεργασία (με ένα κύριο κι ένα εφεδρικό αντίγραφο), όπως φαίνεται στο Σχήμα 7. Το obj1 χρησιμοποιεί ενεργή πλεονασματική επεξεργασία (με δυο αντίγραφα), όπως φαίνεται στο Σχήμα 5. Ο πίνακας 2 συνοψίζει τις τιμές των παραμέτρων που χρησιμοποιήθηκαν για να ρυθμίσουν την κυκλοφορία των αιτήσεων μέσα στο σύστημα και για να ρυθμίσουν την κατανάλωση των πόρων. Η κατανάλωση πόρων εξαρτάται από την ταχύτητα και τον φόρτο των υπολογιστικών σταθμών (hosts) όπου τοποθετούνται τα μέλη των αντικειμένων εξυπηρέτησης.

Αντικείμ. εξυπηρέτησης:	ObjX:SrvRequestAccepting (X=0, 5, 6, 7)	Χωρίς κατάσταση
Αντικείμ. επεξεργασίας:	Obj1, Obj2, Obj3, Obj4	Με κατάσταση
Πολυνηματισμός:	μοντέλο νήμα-ανά-αντικείμενο	
Ανάθεση αιτήσεων εξυπηρέτησης (εξισορρόπηση φόρτου):	Ανά αίτηση διαμοιρασμός φόρτου: εκ περιτροπής (RR) ανάθεση των αιτήσεων στα αντικείμενα ObjX (X=0, 5, 6, 7)	

Πίνακας 1. Η υπολογιστική δομή του συστήματος.

Παράμετροι κυκλοφορίας συστήματος (εκθετικοί με μέσο)										
Κλάση αιτήσεων Class1 (sec)	2.5									
Κλάση αιτήσεων Class2 (sec)	2.5									
Αντίγραφο Αντικειμένου: Παράμετροι κατανάλωσης πόρων	**Αντιγ. 10 Obj.1**	**Αντιγ. 11 Obj.1**	**Αντιγ. 20 Obj.2**	**Αντιγ. 21 Obj.2**	**Αντιγ. 30 Obj.3**	**Αντιγ. 31 Obj.3**	**Αντιγ. 40 Obj.4**	**Αντιγ. 41 Obj.4**	**Αντιγ. X0 Obj.X**	**Αντιγ. X1 Obj.X**
Χρόνοι απόκρισης αιτήσεων Class1 (εκθετικοί με μέσο)	0.52	0.52	0.25 (*)	0.25 (*)	0.7	0.7	0.32	0.32	0.05 (*)	0.05 (*)
Χρόνοι απόκρισης αιτήσεων Class 2 (εκθετικοί με μέσο)	-	-	0.28	0.28	0.7	0.7	-	-	0.05 (*)	0.05 (*)
Χρόνοι Επανεπεξεργασίας ή Επανεκκίνησης αιτήσεων (εκθετικοί με μέσο)	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	-	-
Μεγέθη κατάστασης αντικειμένων (KB)	0.9		0.7		0.5		0.6		-	-
Ταχύτητες αντιγραφής κατάστασης - sec/KB (εκθετικές με μέσο)	0.8	0.8	0.6	0.6	0.6	0.6	0.8	0.8	-	-

Πίνακας 2. Οι παράμετροι της κυκλοφορίας του συστήματος και της κατανάλωσης πόρων.

Το εύρος ζώνης του δικτύου θεωρείται πως είναι άπειρο, ώστε να αποφευχθεί η υπερφόρτωση του μοντέλου με επιπλέον παραμέτρους που δεν σχετίζονται με την απόδοση και την αποτελεσματικότητα της ανοχής σε λάθη. Η επανεξυπηρέτηση των αιτήσεων που είναι αποθηκευμένες στην ουρά καταγραφής, δεν προκαλεί επανεκτέλεση των αντίστοιχων μεθόδων, αλλά απλά την επαναποστολή των ήδη υπολογισμένων απαντήσεων-αποκρίσεων. Η κατανάλωση πόρων κατά την καταγραφή ή αντιγραφή κατάστασης εξαρτάται από τα μεγέθη των καταστάσεων των αντικειμένων, αλλά κι από

την υπολογιστική δυνατότητα που μπορούν να παρέχουν οι υποκείμενοι υπολογιστικοί σταθμοί (state transfer speeds).

Ο πίνακας 3 συνοψίζει τα σχήματα πλεονασματικής επεξεργασίας που χρησιμοποιήθηκαν στις προσομοιώσεις καθώς και τους συνδυασμούς τιμών που αποδόθηκαν στις διάφορες παραμέτρους που καθορίζουν την συμπεριφορά του κάθε σχήματος πλεονασματικής επεξεργασίας.

Τα αποτελέσματα που αποκτήθηκαν κατά την διάρκεια των πειραμάτων προσομοίωσης παρέχουν γνώση για:

- την περίπτωση όπου δεν χρησιμοποιείται καμία πολιτική αναμονής απόκρισης-επαναποστολής. Σε αυτήν την περίπτωση δεν είναι δυνατό να αντιμετωπιστούν λάθη δικτύου ή / και άλλα λάθη (τέτοια όπως αυτά που περιγράφονται στην παράγραφο 4.2).
- τρεις (3) διαφορετικές περιπτώσεις χρήσης αναμονής απόκρισης-επανάκλησης, οι οποίες συνοδεύονται από τα αντίστοιχα επιπλέον κόστη επανεπεξεργασίας-επανάκλησης των αιτήσεων που πιθανώς χάθηκαν.

Σύνθετο σχήμα πλεονασματικής επεξεργασίας	obj0	obj1: classA	obj2: classB	obj3: classC	obj4: classD	obj5	obj6	obj7

Είδος πλεονασματικής επεξεργασίας:	WPR	AR	WPR	WPR	WPR	WPR	WPR	WPR
Παράμετροι Συμπεριφοράς								
Πλήθος Αντιγράφων:	2	2	2	2	2	2	2	2
Διαστήματα καταγραφής/αντιγραφής καταστάσεων (ανά πλήθος αιτήσεων):	Χωρίς κατά-σταση	-	60	30	90	Χωρίς κατά-σταση	Χωρίς κατά-σταση	Χωρίς κατά-σταση
Διαστήματα αναμονής απόκρισης-επανάκλησης (timeouts in sec) ***περίπτωση I***: ***περίπτωση II***: ***περίπτωση III***:	12.0 14.0 16.0	- - -	7.0 9.0 11.0	- - -	- - -	12.0 14.0 16.0	12.0 14.0 16.0	12.0 14.0 16.0

Πίνακας 3. Το σύνθετο Σχήμα Πλεονασματικής Επεξεργασίας που χρησιμοποιήθηκε στις προσομοιώσεις.

Υπολογιστικός κόμβος 1	Αντιγ.00 (obj0)	Αντιγ.11 (obj1)	Αντιγ.51 (obj5)
Υπολογιστικός κόμβος 2	Αντιγ.01 (obj0)		Αντιγ.50 (obj5)
Υπολογιστικός κόμβος 3	Αντιγ.21 (obj2)	Αντιγ.40 (obj4)	
Υπολογιστικός κόμβος 4	Αντιγ.20 (obj2)	Αντιγ.41 (obj4)	
Υπολογιστικός κόμβος 5	Αντιγ.30 (obj3)		
Υπολογιστικός κόμβος 6	Αντιγ.31 (obj3)		
Υπολογιστικός κόμβος 7	Αντιγ.60 (obj6)	Αντιγ.10 (obj1)	Αντιγ.71 (obj7)
Υπολογιστικός κόμβος 8	Αντιγ.61 (obj6)		Αντιγ.70 (obj7)

Πίνακας 4. Η τοποθέτηση των αντιγράφων της κάθε ομάδας αντικειμένων.

Ο πίνακας 4 περιγράφει την τοποθέτηση των μελών των διάφορων ομάδων αντικειμένων στα διαθέσιμα υπολογιστικά κέντρα. Κάθε αντικείμενο εξυπηρέτησης τοποθετείται σε ξεχωριστό υπολογιστικό κέντρο και το μοντέλο πολυνηματισμού που εφαρμόζεται είναι η (όπως φαίνεται και στον πίνακα 1) νήμα-ανα-αντικείμενο (thread-per-object) πολιτική.

Συχνότητα σφαλμάτων :	(r =) 21600 sec									
Αντίγραφα Αντικειμένων:	Αντιγ. X0 Obj.X	Αντιγ. X1 Obj.X	Αντιγ. 10 Obj.1	Αντιγ. 11 Obj.1	Αντιγ. 20 Obj.2	Αντιγ. 21 Obj.2	Αντιγ. 30 Obj.3	Αντιγ. 31 Obj.3	Αντιγ. 40 Obj.4	Αντιγ. 41 Obj.4
Χρόνοι μεταξύ σφαλμάτων (εκθετικοί)	2*r	2*r	2*r	2*r	r	r	r	r	2*r	2*r
Χρόνοι επανεκκίνησης (εκθετικοί)	23.0	23.0	23.0	23.0	23.0	23.0	23.0	23.0	23.0	23.0
Σενάρια Απώλειας κλήσεων (βλ. παραγ. 4.2): ***περίπτωση A***: ***περίπτωση B***: ***περίπτωση C***:	- - -		- - -		- (d) (a)		- (d) (a)		- (d) (a)	
Διαστήματα ελέγχου για λάθη (sec):	2.0, 4.0, 6.0, 8.0, 10.0, 12.0									

Πίνακας 5. Τα μοντέλα λαθών και τα σενάρια απώλειας κλήσεων που χρησιμοποιήθηκαν στις προσομοιώσεις.

Τελικά ο Πίνακας 5 δείχνει τα μοντέλα των λαθών και τα σενάρια απώλειας κλήσεων που χρησιμοποιήθηκαν στην προσομοίωση. Έγινε χρήση παραμετρικών μοντέλων λαθών (δηλαδή που η εμφάνισή τους εξαρτάται από κάποια/ες παραμέτρους), αφού τα λάθη είναι εξ' ορισμού σπάνια γεγονότα και είναι πάντα χρήσιμο, αν όχι απαραίτητο, η αναφορά της ευαισθησίας των παραγόμενων αποτελεσμάτων να γίνεται σε σχέση με την συχνότητα εμφάνισης των λαθών που κάθε φορά υποτίθεται. Ο πίνακας καθορίζει τα διαστήματα μεταξύ των λαθών, τους χρόνους επανάκτησης καθώς και τα τρία (3) διαφορετικά σενάρια λαθών που χρησιμοποιήθηκαν. Οι μηχανισμοί ανίχνευσης των λαθών που χρησιμοποιήθηκαν περιγράφονται από τα έξι (6) διαφορετικά διαστήματα παρακολούθησης που φαίνονται στον ίδιο πίνακα. Στην υλοποίηση κάθε περίπτωση συνοδεύεται από τα αντίστοιχα κόστη που επιφέρει στο σύστημα και που περιγράφονται στην παράγραφο 5.4.

7.2 Τα Αποτελέσματα της Απόδοσης και της Αποτελεσματικότητας της Ανοχής σε Λάθη

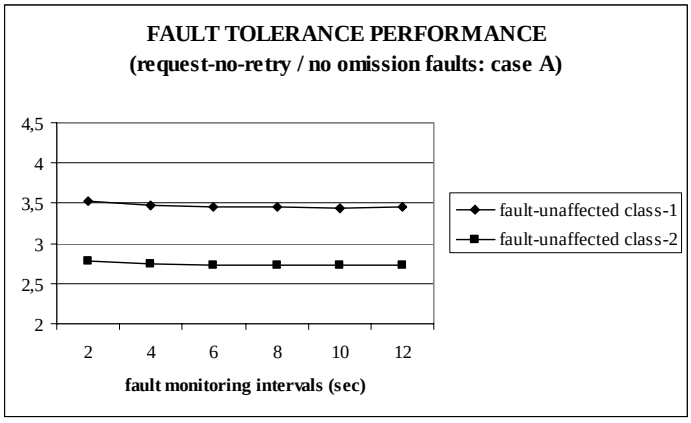
Στο άρθρο [11], στο οποίο έχει ήδη γίνει αναφορά, βρέθηκε ότι *η απόδοση κι η αποτελεσματικότητα της ανοχής σε λάθη εξαρτάται από τον φόρτο του συστήματος* (κατανομές άφιξης των αιτήσεων). Το να ληφθεί υπόψη ο προβλεπόμενος φόρτος του συστήματος είναι απόλυτη ανάγκη: μπορεί να αποτελέσει τον αποφασιστικό παράγοντα στην απόφαση για τον τρόπο δόμησης των σχημάτων πλεονασματικής επεξεργασίας ώστε να ικανοποιούνται τα απαιτούμενα επίπεδα ποιότητας επεξεργασίας. Επιπρόσθετα, ένα κατάλληλα επιλεγμένο σχήμα πλεονασματικής επεξεργασίας επίσης εξασφαλίζει προοπτικές λήψης αποφάσεων όσον αφορά τις αντισταθμίσεις μεταξύ απόδοσης κι αποτελεσματικότητας της ανοχής σε λάθη, σε ένα περιβάλλον συνεχώς μεταβαλλόμενων αναγκών ποιότητας εξυπηρέτησης.

Όπως έχει ήδη αναφερθεί, η παρούσα εργασία δεν παρουσιάζει κάποια συγκεκριμένη συστηματική μέθοδο σχεδιασμού ποιότητας εξυπηρέτησης. Η εργασία επικεντρώνεται κυρίως στο να παρουσιάσει μια προσέγγιση αποτίμησης ως ένα γενικής χρήσης εργαλείο που μπορεί να χρησιμοποιηθεί για την μελέτη μιας σειράς σημαντικών παραγόντων αντιστάθμισης όπως:

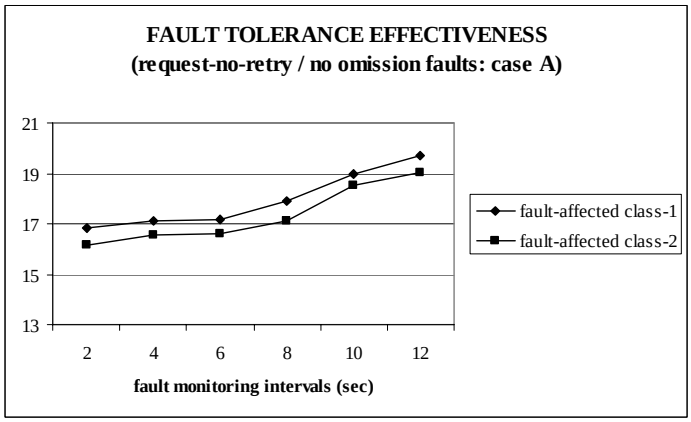
- την επιλογή κατάλληλων σημείων καταγραφής κατάστασης κι αντιγραφής κατάστασης για τα αντικείμενα που χρησιμοποιούν παθητική πλεονασματική επεξεργασία (με αυτό το πρόβλημα ασχολείται το [11]).
- και την ανάγκη αποφυγής υπερβολικά συχνών διαστημάτων αναμονής απόκρισης-επανάκλησης, γιατί προκαλούν υψηλά επιπλέον κόστη και δεν βελτιώνουν την αποτελεσματικότητα της ανοχής σε λάθη, όπως ίσως θα περίμενε κανείς (τα αποτελέσματα που παρουσιάζονται στην συνέχεια βοηθούν στην καλύτερη κατανόηση αυτού του προβλήματος).

Η υβριδική προσομοίωση αξιοπιστίας και κυκλοφορίας συστήματος μαζί, σε συνεργασία με την παρουσιαζόμενη προσέγγιση αποτίμησης μπορούν να αποτελέσουν το βασικό στοιχείο για την λήψη απόφασης σχετικά με οποιονδήποτε άλλον παράγοντα αντιστάθμισης ή ανάθεσης κατάλληλων τιμών σε σύνθετους συνδυασμούς παραμέτρων συμπεριφοράς (π.χ. σημεία καταγραφής κατάστασης, διαστήματα αντιγραφής κατάστασης, διαστήματα αναμονής απόκρισης-επανάκλησης κτλ.).

Τα παρακάτω αποτελέσματα προέκυψαν μετά από την προσομοίωση του συστήματος που περιγράφηκε νωρίτερα. Στις προσομοιώσεις χρησιμοποιήθηκαν οι κατανομές άφιξης των αιτήσεων που ορίστηκαν στον πίνακα 2. Τα αποτελέσματα των γραφικών παραστάσεων παράχθηκαν με 5% διάστημα εμπιστοσύνης και εύρος όχι περισσότερο από 3% της εκάστοτε εκτιμώμενης τιμής.



(a)



(b)

Σχήμα 11. Η απόδοση κι η αποτελεσματικότητα της ανοχής σε λάθη για το σχήμα του πίνακα 3 που δεν χρησιμοποιεί καμία πολιτική αναμονής απόκρισης – επανάκλησης.

Στο σχήμα 11 παρουσιάζονται η παρατηρούμενη απόδοση (μέσοι όροι των χρόνων απόκρισης των αιτήσεων που δεν επηρεάστηκαν από λάθη) κι η αποτελεσματικότητα (μέσοι όροι των χρόνων απόκρισης των αιτήσεων που επηρεάστηκαν από λάθη) της ανοχής σε λάθη για το σύνθετο σχήμα πλεονασματικής επεξεργασίας του πίνακα 3, στην περίπτωση όμως όταν δεν χρησιμοποιείται καμία πολιτική αναμονής απόκρισης-

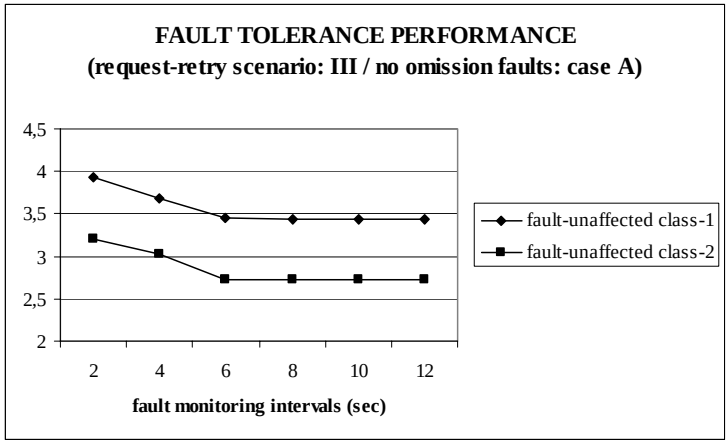
επανάκλησης για την αντιμετώπιση των πιθανών λαθών με τα οποία ασχολείται η εργασία και που περιγράφονται στην παράγραφο 4.2

Παρατηρείται μια αξιοπρόσεχτη βελτίωση της αποτελεσματικότητας (σχήμα 11b), όταν μειώνεται το διάστημα παρακολούθησης από 12 σε 6 sec. Παρόλα αυτά, δεν παρατηρείται σημαντική επιπλέον βελτίωση όταν το διάστημα μειώνεται ακόμη περισσότερο. Η απόδοση της ανοχής σε λάθη (σχήμα 11a) αναφέρεται στην τεράστια πλειοψηφία των διακινούμενων αιτήσεων και δεν επηρεάζεται ιδιαίτερα από το κόστος που επιβάλλεται από τον εφαρμοζόμενο μηχανισμό ανίχνευσης των λαθών.

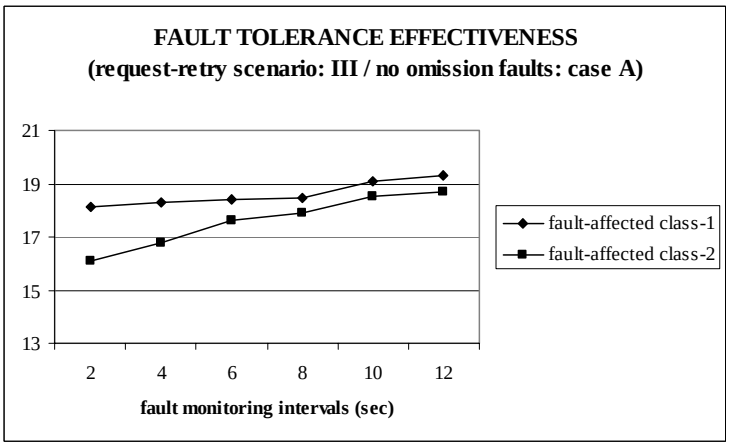
Από την άλλη όμως, ένας σχεδιαστής συστήματος θα προτιμούσε να χρησιμοποιήσει κάποια, έστω, πολιτική αναμονής απόκρισης-επανάκλησης, ώστε να παρέχεται στο σύστημά του η δυνατότητα αντιμετώπισης των πιθανών λαθών. Το επιπλέον κόστος για την επανάκληση των αιτήσεων που είναι πιθανώς χαμένες έχει σαν αποτέλεσμα χειρότερη απόδοση (σχήματα 12a, 12c, 12e) και χειρότερη αποτελεσματικότητα (σχήματα 12b, 12d, 12f) σε σχέση με την περίπτωση του σχήματος 11.

Το σενάριο αναμονής απόκρισης-επανάληψης που ορίζεται ως περίπτωση III στον πίνακα 3 (σχήματα 12a, 12b) επιτρέπει την χρήση διαστημάτων παρακολούθησης λαθών από 6 ως 12 sec., χωρίς σημαντικά επιπλέον κόστη. Ωστόσο, πιο μικρά διαστήματα ανίχνευσης λαθών επιβαρύνουν την απόδοση της ανοχής σε λάθη (σχήμα 12a) με ιδιαίτερα υψηλά κόστη.

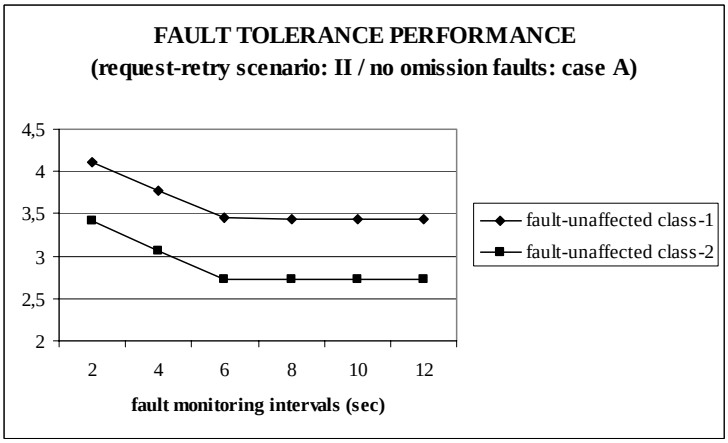
Όταν τα διαστήματα αναμονής απόκρισης-επανάκλησης που χρησιμοποιούν τα αντικείμενα γίνονται ακόμη πιο συχνά (περιπτώσεις σεναρίων II και I στον πίνακα 3) παρατηρούμε τα ίδια ή και μεγαλύτερα ακόμη επιπλέον κόστη (σχήματα 12c και 12e αντίστοιχα) και χειρότερη αποτελεσματικότητα ανοχής σε λάθη (σχήματα 12d και 12f αντίστοιχα). Η περίπτωση του σεναρίου I διαστήματος αναμονής απόκρισης-επανάκλησης δεν είναι καθόλου αποτελεσματική (χρήση υπερβολικά συχνών διαστημάτων).


FAULT TOLERANCE PERFORMANCE
(request-retry scenario: III / no omission faults: case A)
4,5
4
3,5
3
2,5
2
2
4
6
8
10
12
fault monitoring intervals (sec)
fault-unaffected class-1
fault-unaffected class-2

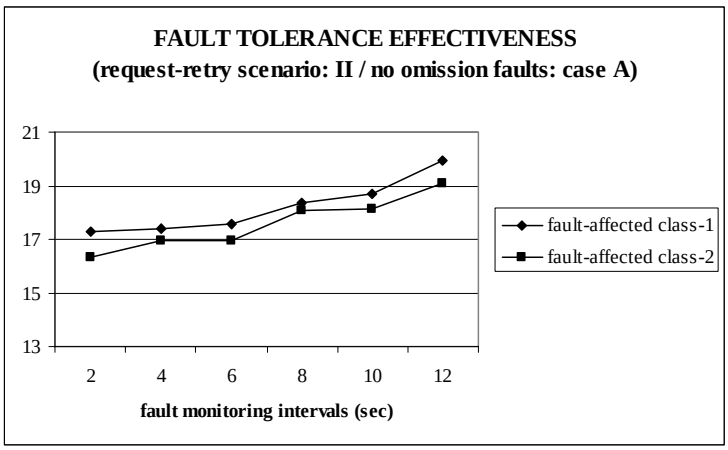
(a)


FAULT TOLERANCE EFFECTIVENESS
(request-retry scenario: III / no omission faults: case A)
21
19
17
15
13
2
4
6
8
10
12
fault monitoring intervals (sec)
fault-affected class-1
fault-affected class-2

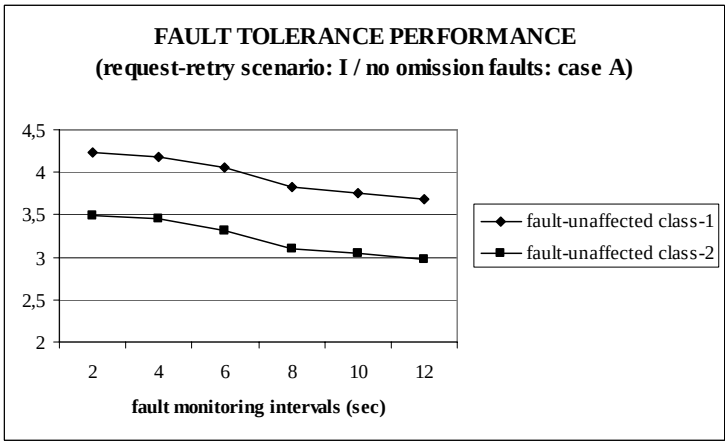
(b)


FAULT TOLERANCE PERFORMANCE
(request-retry scenario: II / no omission faults: case A)
4,5
4
3,5
3
2,5
2
2
4
6
8
10
12
fault monitoring intervals (sec)
fault-unaffected class-1
fault-unaffected class-2

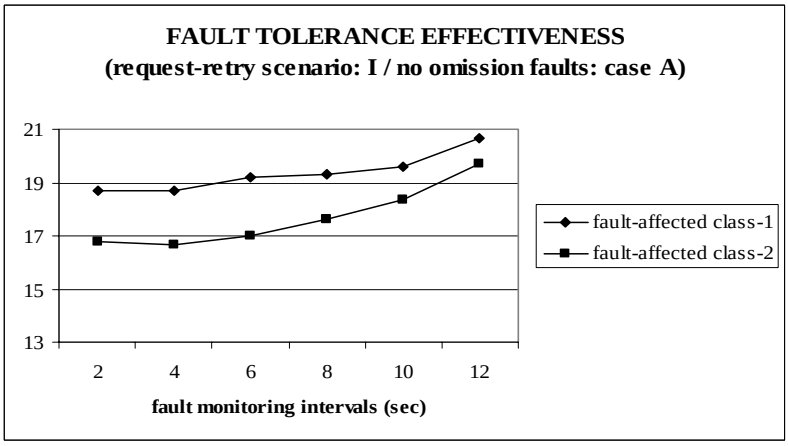
(c)



(d)



(e)



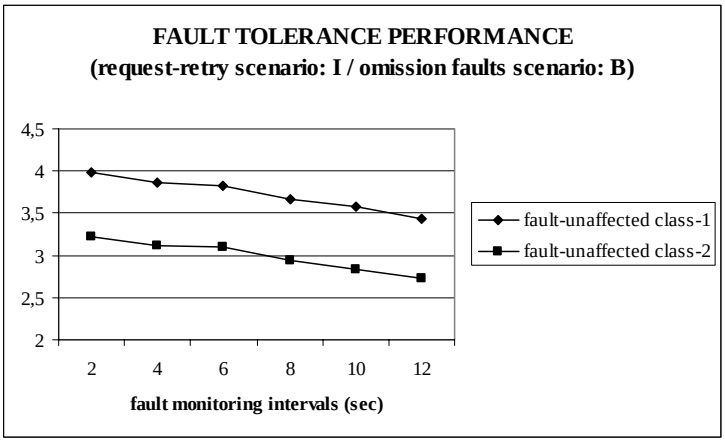
(f)

Σχήμα 12. Η απόδοση κι η αποτελεσματικότητα της ανοχής σε λάθη για τα σχήματα πλεονασματικής επεξεργασίας του πίνακα 3 με διαφορετικές πολιτικές αναμονής απόκρισης – επανάκλησης.

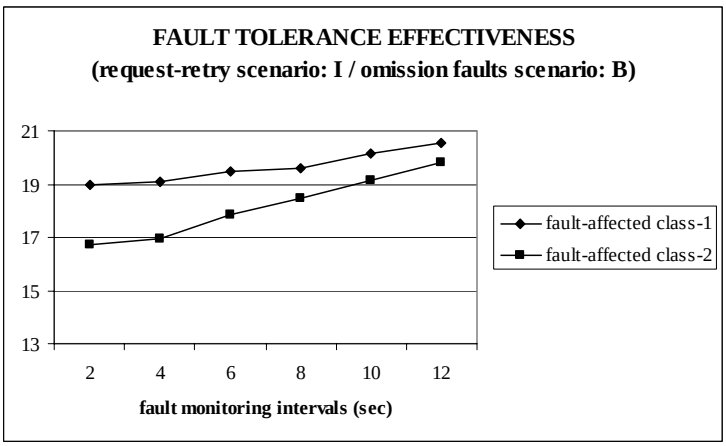
Οι προσομοιώσεις επίσης υποδεικνύουν το πώς η απόδοση κι η αποτελεσματικότητα επηρεάζονται από τα διαφορετικά σενάρια απώλειας κλήσεων των λαθών (σχήμα 13).

Τα σχήματα 13a και 13b αναφέρονται στο σενάριο απώλειας B του πίνακα 5. Δηλαδή όταν σε ένα παθητικής πλεονασματικής επεξεργασίας αντικείμενο συμβαίνει λάθος, τα ήδη αποθηκευμένα αιτήματα στην ουρά χάνονται και τα αιτήματα που καταφθάνουν κατά την διάρκεια που το αντικείμενο δεν είναι διαθέσιμο, επίσης χάνονται (στα ενεργητικής πλεονασματικής επεξεργασίας αντικείμενα χρησιμοποιείται σε όλες τις περιπτώσεις αυτό το σενάριο λάθους). Παρατηρούνται μικρές βελτιώσεις στην απόδοση (σε σχέση με το 12e) οι οποίες οφείλονται στις άδειες ουρές που βρίσκουν οι μη επηρεασμένες από λάθη αιτήσεις όταν καταφθάνουν σε ένα κύριο αντίγραφο ομάδας αντικειμένων που μόλις ολοκλήρωσε την διαδικασία της επανάκτησης.

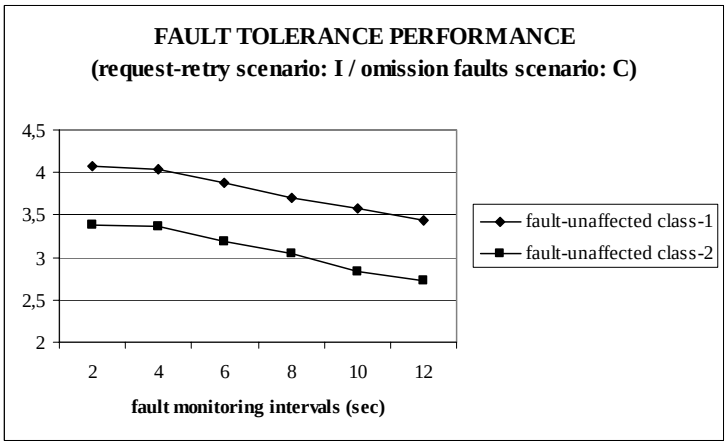
Τα σχήματα 13c και 13d αναφέρονται στο σενάριο απώλειας C του πίνακα 5. Δηλαδή όταν σε ένα παθητικής πλεονασματικής επεξεργασίας αντικείμενο συμβαίνει λάθος, κανένα από τα ήδη αποθηκευμένα αιτήματα στην ουρά δεν χάνεται, αλλά τα αιτήματα που καταφθάνουν κατά την διάρκεια που το αντικείμενο δεν είναι διαθέσιμο, χάνονται. Η απόδοση (σχήμα 13c) της ανοχής σε λάθη πάλι βελτιώνεται σε σχέση με την περίπτωση όπου καμία αίτηση δεν χάνεται (σενάριο A στον πίνακα 5 και σχήμα 12e), αλλά σε σχέση με την περίπτωση του σχήματος 13a η απόδοση είναι λίγο χειρότερη.



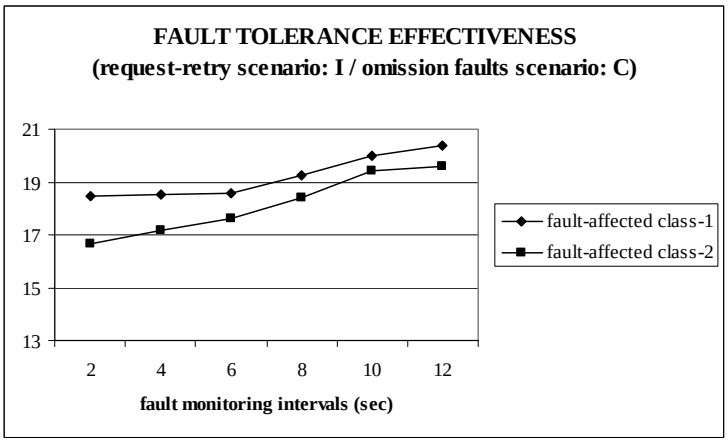
(a)



(b)



(c)



(d)

Σχήμα 13. Η απόδοση κι η αποτελεσματικότητα της ανοχής σε λάθη κάτω από διαφορετικά σενάρια απώλειας κλήσεων (Πίνακας 5)

Διαφορετικά σενάρια απώλειας κλήσεων απαιτούν από τα αντικείμενα διαφορετική αντιμετώπιση για την επίτευξη ανοχής στα λάθη (π.χ. [2], [5], [9], [16], [19], [25]).

Αναμένονται σημαντικές διαφορές για μεγαλύτερα επίπεδα φόρτου συστήματος κι αυτό πρέπει να λαμβάνεται υπόψη στον σχεδιασμό των συστημάτων που χρησιμοποιούν πλεονασματική επεξεργασία και πολιτικές αναμονής απόκρισης-επανάκλησης με στόχο να ικανοποιήσουν συγκεκριμένα επίπεδα ποιότητας εξυπηρέτησης.

ΚΕΦΑΛΑΙΟ 8ο

Συμπεράσματα

Σε αυτήν την εργασία παρουσιάστηκε μια ποσοτική προσέγγιση αποτίμησης για εφαρμογές αξιόπιστης εξυπηρέτησης οι οποίες πρέπει να ικανοποιούν κάποια, πιθανώς προσυμφωνημένα επίπεδα ποιότητας εξυπηρέτησης ή / και εγγυήσεις χρόνων εξυπηρέτησης. Σε σύγκριση με άλλες προσεγγίσεις αποτίμησης αξιοπιστίας, η εργασία παρουσιάζει μια υβριδική *αξιοπιστίας* ***και*** *κυκλοφορίας συστήματος* προσομοίωση κι επικεντρώνεται σε μετρικές επί των χρόνων απόκρισης, οι οποίες ποσοτικοποιούν ξεχωριστά την αποτελεσματικότητα και την απόδοση της ανοχής σε λάθη.

Το λογισμικό της προσομοίωσης αναπτύχθηκε κατά την διάρκεια της εργασίας σε γλώσσα προγραμματισμού C++. Στην υλοποίηση συμπεριλαμβάνονται όλοι οι παράγοντες και οι παράμετροι που έχουν αναφερθεί στα προηγούμενα κεφάλαια. Επίσης, τα αποτελέσματα παράχθηκαν και η παρουσίασή τους γίνεται σύμφωνα με την προσέγγιση αποτίμησης που παρουσιάζει η εργασία.

Η προσέγγιση αποτίμησης που παρουσιάζεται ανοίγει τις ακόλουθες προοπτικές:

- να λαμβάνει υπόψη σύνθετες αλληλεπιδράσεις οι οποίες αλλιώτικα θα αποδίδονταν ξεχωριστά σε διαφορετικές επιλογές σχεδίασης (ανοχή σε λάθη, εξισορρόπηση φόρτου, πολυνηματισμός, κτλ.),
- να αιχμαλωτίζει την ουσία των πιο σημαντικών παραγόντων αντιστάθμισης της ανοχής σε λάθη,
- να υποστηρίζει συνδυασμένη λήψη απόφασης σχετικά με την παραμετροποίηση διαφορετικών ιδιοτήτων, όπως σχετικά με τα σημεία καταγραφής κατάστασης, τα διαστήματα αντιγραφής κατάστασης, τα διαστήματα αναμονής απόκρισης-επανάκλησης και άλλα,
- να παρέχει εκτιμήσεις για υποψήφια επίπεδα ποιότητας εξυπηρέτησης, που συχνά προσδιορίζονται μετά από συζήτηση μεταξύ των παροχέων της εξυπηρέτησης και των πελατών,

- να εξερευνάει τις προοπτικές ενσωμάτωσης ενός σχήματος πλεονασματικής επεξεργασίας σε ένα περιβάλλον όπου συνεχώς μεταβάλλονται οι ανάγκες ποιότητας εξυπηρέτησης.

Η προσέγγιση αποτίμησης που παρουσιάζει η εργασία μπορεί επίσης να αποτελέσει θεμέλιο λίθο για την ανάπτυξη συστηματικών μεθόδων σχεδιασμού υψηλής ποιότητα εξυπηρέτησης (QoS design methods) στις οποίες:

- τα υποψήφια σχήματα πλεονασματικής επεξεργασίας θα μπορούν να συγκρίνονται σύμφωνα με τις συνθήκες της βέλτιστης αποτελεσματικότητάς τους (που είναι το μοναδικό κριτήριο που μπορεί να κάνει δυνατή μια τέτοια σύγκριση), κι επίσης
- θα είναι δυνατό να προσδιοριστούν εκείνες οι τιμές που πρέπει να αποδοθούν στις ιδιότητες συμπεριφοράς των, οι οποίες θα επιφέρουν το χαμηλότερο δυνατό κόστος, αλλά και θα ικανοποιούν τον όποιο συγκεκριμένο στόχο επίτευξης κατάλληλων επιπέδων ποιότητας εξυπηρέτησης.

Η υβριδική αξιοπιστίας και κυκλοφορίας συστήματος προσομοίωση και η προσέγγιση αποτίμησης που παρουσιάστηκαν σε αυτήν την εργασία αποτελούν ένα πολύτιμο και γενικής χρήσης εργαλείο που μπορεί να χρησιμοποιηθεί για την μελέτη της απόδοσης και της αποτελεσματικότητας της ανοχής σε λάθη μέσα σε πολλά και πολύ διαφορετικά περιεχόμενα (πχ. αλγόριθμοι συνδυασμένης καταγραφής κατάστασης και αποθήκευσης των εξυπηρετημένων αιτήσεων στην ουρά καταγραφής, ανοχή σε λάθη βασισμένη σε συναλλαγές, όπως στο [3] και [6] κτλ.). Τελικά, η παρουσιαζόμενη προσέγγιση μπορεί επίσης να αποτελέσει βασικό στοιχείο μοντέλων (που βασίζονται σε UML) απόδοσης αξιόπιστων εφαρμογών.

ΓΛΩΣΣΑΡΙ

Ανοχή σε σφάλματα (fault tolerance): το σύστημα να παρέχει μηχανισμούς, τέτοιους ώστε εάν συμβεί σφάλμα σε κάποιο από τα πλεονασματικά αντικείμενα ενός αντικειμένου, αυτό να μπορεί να διορθωθεί έτσι ώστε το πλεονασματικό αντικείμενο να συνεχίσει κανονικά τη λειτουργία του.

Αξιοπιστία (reliability): όταν σε ένα σύστημα υπάρχει κατανομή του φόρτου εργασίας σε πολλούς εξυπηρέτες κι έτσι η διακοπή λειτουργίας ενός εξυπηρέτη δε θα προκαλέσει και την κατάρρευση του συστήματος ως ολότητα.

Διαδιδόμενα σφάλματα (commission faults): Αυτά τα σφάλματα πρέπει να ανιχνεύονται από την ενεργητική πλεονασματική επεξεργασία (active replication).

συμβαίνουν όταν ένα αντικείμενο δημιουργεί εσφαλμένα αποτελέσματα.

Ενεργητική πλεονασματική επεξεργασία (active replication): είδος πλεονασματικής επεξεργασίας στο οποίο δεν υπάρχει διαχωρισμός των αντιγράφων σε κύρια και εφεδρικά γιατί όλα τα αντίγραφα του αντικειμένου εκτελούν όλες τις λειτουργίες με την

ίδια σειρά, ανεξάρτητα όμως το ένα από	το άλλο.
Επαναφορά (recovery):	η διαδικασία κατά την οποία γίνεται διόρθωση του σφάλματος που συνέβη σε κάποιο πλεονασματικό αντικείμενο.
Θερμή παθητική πλεονασματική επεξεργασία (warm passive replication):	είδος πλεονασματικής επεξεργασίας στο οποίο μόνο ένα από τα αντίγραφα, το χαρακτηριζόμενο ως κύριο, διεκπεραιώνει όλες τις λειτουργίες εφόσον βρίσκεται στην κανονική κατάσταση λειτουργίας του. Τα υπόλοιπα αντίγραφα συγχρονίζονται με το κύριο ανά περιοδικά χρονικά διαστήματα.
Ισχυρή συνέπεια ομάδας αντικειμένων (strong replica consistency):	όταν στην περίπτωση της ενεργητικής πλεονασματικής επεξεργασίας (active replication), όλα τα αντίγραφα που έφεραν εις πέρας την ίδια λειτουργία, έχουν την ίδια κατάσταση.
Μοντέλο πελάτη-εξυπηρέτη (client-server):	το σύστημα αποτελείται από ένα σύνολο αντικειμένων, τα οποία αλληλεπιδρούν μεταξύ τους με σκοπό να εξυπηρετήσουν ένα σύνολο από κλήσεις (requests). Τα αντικείμενα που εξυπηρετούν τις κλήσεις ονομάζονται εξυπηρέτες (server objects) και τα αντικείμενα που στέλνουν κλήσεις προς

εξυπηρέτηση ονομάζονται πελάτες (client objects).

Μοντέλο σφάλμα-παύση (fail-stop model):

όταν συμβαίνουν σφάλματα τα αντικείμενα σταματούν πλήρως οποιαδήποτε λειτουργία και καταρρέουν εντελώς, χωρίς να στέλνουν ψεύτικα (spurious) μηνύματα.

Ομάδα αντικειμένων (object group):

το σύνολο των πλεονασματικών αντικειμένων που συνιστούν ένα αντικείμενο (είτε αυτό είναι εξυπηρέτης είτε πελάτης).

Ουρά καταγραφής (message log):

μια εγγραφή που περιέχει μηνύματα και καταστάσεις αντικειμένων και δημιουργείται για να διασφαλίσει ότι είναι δυνατή η επαναφορά ενός αντικειμένου μετά από ένα σφάλμα.

Πλεονασματικά αντικείμενα (replicas):

κάθε αντικείμενο (είτε αυτό είναι εξυπηρέτης είτε πελάτης), διαθέτει αντίγραφα του εαυτού του και τα αντίγραφα αυτά ονομάζονται πλεονασματικά αντικείμενα.

Σφάλμα (fault):

όταν η συμπεριφορά ενός εξαρτήματος του συστήματος (π.χ. ενός αντικειμένου) προκαλεί λαθεμένη

συμπεριφορά σε όλο το σύστημα. Είναι η επέκταση σε όλο το σύστημα, της εκδήλωσης ενός σφάλματος σε ένα εξάρτημα.

Ψυχρή παθητική πλεονασματική επεξεργασία (cold passive replication):

είδος πλεονασματικής επεξεργασίας στο οποίο είναι το κύριο αντίγραφο μόνο που διεκπεραιώνει όλες τις λειτουργίες. Τα υπόλοιπα αντίγραφα μένουν αδρανή για όσο διάστημα το κύριο αντίγραφο λειτουργεί κανονικά. Όταν συμβεί σφάλμα, ένα από τα εφεδρικά αντίγραφα συγχρονίζεται με το κύριο και συνεχίζει τις λειτουργίες του.

Ακολουθεί (μια εναλλακτική) περιγραφή του διαγράμματος μεταβάσεων κατάστασης αντίγραφου ενεργητικής πλεονασματικής επεξεργασίας που υιοθετήθηκε στην υλοποίηση της εργασίας:

Οι δυνατές καταστάσεις στις οποίες μπορεί να βρεθεί ένα αντίγραφο στην περίπτωση της ενεργητικής πλεονασματικής επεξεργασίας είναι: α) κανονική κατάσταση (normal), β) κατάσταση σφάλματος (fault), γ) κατάσταση επαναφοράς (recovering), δ) κατάσταση αναμονής αντιγραφής κατάστασης (st_wait) και ε) κατάσταση μεταφοράς/αντιγραφής κατάστασης (state transferring). Ένα αντίγραφο βρίσκεται στην κανονική κατάσταση λειτουργίας του όταν δεν του συμβαίνει κανένα σφάλμα. Αντιθέτως, όταν του συμβεί σφάλμα μεταβαίνει στην κατάσταση σφάλματος. Όταν ανιχνευθεί το σφάλμα, το αντίγραφο μπορεί να μεταβεί μόνο στην κατάσταση επαναφοράς κατά τη διάρκεια της οποίας διορθώνεται το σφάλμα που συνέβη. Από εκεί εάν ξανασυμβεί σφάλμα επιστρέφει στην κατάσταση σφάλματος. Εάν όμως η διαδικασία επαναφοράς ολοκληρωθεί, γίνεται επανεκκίνηση του αντιγράφου και μεταβαίνει στην κατάσταση αναμονής κατά την οποία περιμένει να περατώσει κάποιο από τα υπόλοιπα αντίγραφα τις λειτουργίες του ώστε να συγχρονιστεί μαζί του. Από την κατάσταση αναμονής μπορεί είτε να μεταβεί στην κατάσταση σφάλματος, εάν

συμβεί πάλι σφάλμα, είτε να μεταβεί στην κατάσταση μεταφοράς/αντιγραφής κατάστασης. Τέλος από εκεί υπάρχει η πιθανότητα να γυρίσει πάλι στην κατάσταση σφάλματος, εάν φυσικά ξανασυμβεί σφάλμα κατά τη διάρκεια μεταφοράς/αντιγραφής κατάστασης, ή να επιστρέψει στην κανονική κατάσταση λειτουργίας του.

Ακολουθεί (μια εναλλακτική) περιγραφή του διαγράμματος μεταβάσεων κατάστασης αντίγραφου θερμής παθητικής πλεονασματικής επεξεργασίας που υιοθετήθηκε στην υλοποίηση της εργασίας:

Όσον αφορά τις μεταβάσεις στο κύριο αντίγραφο ισχύουν τα ακόλουθα: το αντίγραφο βρίσκεται στην κανονική κατάσταση λειτουργίας του όταν δεν του συμβαίνει κανένα σφάλμα. Αντιθέτως, όταν του συμβεί σφάλμα, μεταβαίνει στην κατάσταση σφάλματος. Όταν ανιχνευθεί το σφάλμα, γίνεται επανεκκίνηση του αντιγράφου είτε ως κύριο είτε ως εφεδρικό. Το αντίγραφο επανεκκινείται ως εφεδρικό εάν βρεθεί έστω ένα άλλο εφεδρικό αντίγραφο, το οποίο να λειτουργεί κανονικά για να το αντικαταστήσει ως κύριο. Εάν γίνει επανεκκίνηση του αντιγράφου ως εφεδρικό τότε αλλάζει η συμπεριφορά του και από το σχήμα 5(α) μεταβαίνει στο σχήμα 5(β) και συνεχίζει από το σημείο «επανεκκίνηση ως εφεδρικό» (“restart as backup”). Εάν όμως δε βρεθεί κανένα εφεδρικό για να αντικαταστήσει το κύριο, τότε το αντίγραφο επανεκκινείται ως κύριο και μεταβαίνει κατευθείαν στην κατάσταση επαναφοράς κατά τη διάρκεια της οποίας διορθώνεται το σφάλμα που συνέβη. Από εκεί εάν ξανασυμβεί σφάλμα επιστρέφει στην κατάσταση σφάλματος. Εάν όμως η διαδικασία επαναφοράς ολοκληρωθεί, τότε το κύριο αντίγραφο θα επιστρέψει στην κανονική κατάσταση λειτουργίας του, αφού προηγουμένως ανακτήσει από την ουρά καταγραφής (message log) την κατάσταση που είχε πριν του συμβεί το σφάλμα. Από την κανονική κατάσταση λειτουργίας το αντίγραφο μπορεί επίσης να μεταβεί στην κατάσταση μεταφοράς/αντιγραφής κατάστασης όποτε συμβαίνει αντιγραφή κατάστασης. Από εκεί μπορεί να επιστρέψει στην κανονική κατάσταση όταν περατωθεί η διαδικασία ή να μεταβεί στην κατάσταση σφάλματος εάν φυσικά συμβεί σφάλμα.

Όσον αφορά τις μεταβάσεις στα εφεδρικά αντίγραφα ισχύουν τα εξής: το αντίγραφο βρίσκεται στην κανονική κατάσταση λειτουργίας του όταν δεν του συμβαίνει κανένα σφάλμα. Εάν μετατραπεί σε κύριο αντίγραφο, για να αντικαταστήσει όπως έχουμε πει το προηγούμενο κύριο αντίγραφο, τότε αλλάζει συμπεριφορά και μεταβαίνει στο σχήμα

5(α) συνεχίζοντας τη λειτουργία του από το σημείο «επανεκκίνηση ως κύριο» (“become the primary”). Όταν βρίσκεται στην κανονική κατάσταση και του συμβεί σφάλμα, μεταβαίνει στην κατάσταση σφάλματος. Όταν ανιχνευθεί το σφάλμα, γίνεται επανεκκίνηση του αντιγράφου ως εφεδρικό και έπειτα μεταβαίνει στην κατάσταση επαναφοράς κατά τη διάρκεια της οποίας διορθώνεται το σφάλμα που συνέβη. Από εκεί εάν ξανασυμβεί σφάλμα επιστρέφει στην κατάσταση σφάλματος. Εάν όμως η διαδικασία επαναφοράς ολοκληρωθεί, τότε το εφεδρικό αντίγραφο μεταβαίνει στην κατάσταση μεταφοράς/αντιγραφής κατάστασης. Από εκεί υπάρχει η πιθανότητα να γυρίσει πάλι στην κατάσταση σφάλματος, εάν φυσικά ξανασυμβεί σφάλμα κατά τη διάρκεια μεταφοράς/αντιγραφής κατάστασης, ή να επιστρέψει στην κανονική κατάσταση λειτουργίας του. Τέλος από την κανονική κατάσταση λειτουργίας το αντίγραφο μπορεί επίσης να μεταβεί στην κατάσταση μεταφοράς/αντιγραφής κατάστασης εάν συμβεί αντιγραφή κατάστασης.

ΑΝΑΦΟΡΕΣ

Α. Βιβλία

1. «Σύγχρονα Λειτουργικά Συστήματα», A. S. Tanenbaum, Τόμος Β' (Κατανεμημένα Συστήματα), Εκδόσεις Παπασωτηρίου (1994)
2. «Συστήματα Παράλληλης Επεξεργασίας», Γ.Κ. Παπακωνσταντίνου, Θ.Α. Θεοχάρης, Π.Δ. Τσανάκας, Εκδόσεις Συμμετρία (1994)
3. "Distributed Operating Systems, The Logical Design", A. Goscinski, Addison-Wesley Publishing Company (1991)

Β. Δικτυακοί Τόποι

1. http://www.math.hkbu.edu.hk/UniformDesign/, 2000, Uniform Design web pages
2. http://www.wikipedia.com , Λεξικό ορολογιών
3. http://wos.ekt.gr, Citation Databases
4. http://ipac.lib.auth.gr, online κατάλογος του Α.Π.Θ.

Γ. Άρθρα

1. C. Basile, K. Whisnant, Z. Kalbarczyk, R. Iyer, Loose synchronization of multi-threaded replicas, Proceedings of the 21st IEEE Symposium on Reliable Distributed Systems (SRDS 02), IEEE Computer Society Press, Osaka University, Suita, Japan, 2002, pp. 250-255
2. T. Bennani, L. Blain, L. Courtes, J.-C. Fabre, M.-O. Killijian, E. Marsden, F. Taiani, Implementing simple replication protocols using CORBA portable interceptors and Java serialization, Proceedings of the IEEE/IFIP International Conference on Dependable Systems and Networks (DSN 04), IEEE Computer Society Press, Florence, Italy, 2004, pp. 549-554

3. I. Crnkovic, M. Larsson, Classification of quality attributes for predictability in component based systems, Supplemental Volume of the 2004 International Conference on Dependable Systems and Networks, Florence, Italy, 2004, pp. 307-311
4. O. Das, C. M. Woodside, The fault-tolerant layered queuing network model for performability of distributed systems, Proceedings of the IEEE International Computer Performance and Dependability Symposium (IPDS 98), Durham, North Carolina, pp. 132-141, 1998
5. P. Felber, R. Guerraoui, A. Schiper, Replication of CORBA Objects, Distributed Systems, Lecture Notes in Computer Science 1752, Springer Verlag, 2000, pp. 254-276
6. J. Fraga, F. Siqueira, F. Favarim, An adaptive fault-tolerant component model, Proceedings of the Ninth IEEE International Workshop on Object-Oriented Real-Time Dependable Systems (WORDS'03), IEEE Computer Society, Capri Island, Italy, 2003, pp. 179-186
7. S. Garg, Y. Huang, C. M. R. Kintala, K. S. Trivedi, S. Yajnik, Performance and reliability evaluation of passive replication schemes in application level fault tolerance, Proceedings of the 29th Annual International Symposium on Fault-Tolerant Computing, IEEE, Madison, Wisconsin, USA, 1999, pp. 322-329
8. K. K. Goswami, R. K. Iyer, L. Young, DEPEND: A simulation-based environment for system level dependability analysis, IEEE Transactions on Computers, 46, 1, 1997, pp. 60-74
9. R. Guerraoui, P. Eugster, P. Felber, B. Garbinato, K. Mazouni, Experiences with object group systems, Software: Practice & Experience, 30, 12, 2000, pp. 1375-1404
10. E. J. Henley, H. Kumamoto, Reliability engineering and risk assessment, Prentice-Hall, 1981
11. P. Katsaros, C. Lazos, Optimal object state transfer - recovery policies for fault tolerant distributed systems, Proceedings of the IEEE/IFIP International Conference on Dependable Systems and Networks (DSN 04), IEEE Computer Society, Florence, Italy, 2004, pp. 762-771
12. P. Katsaros, E. Angelis, C. Lazos, Applied multiresponse metamodeling for queuing network simulation experiments: problems and perspectives, Proceedings of the 4th EUROSIM Congress on Modeling and Simulation, EUROSIM, Delfts, The Netherlands, 2001
13. A. M. Law, J. C. Carson, A sequential procedure for determining the length of a steady state simulation, Operations Research, Vol. 27, 1979, pp. 1011-1025
14. M. Lindermeier, Load management for distributed object-oriented environments, International Symposium on Distributed Objects and Applications (DOA'00), IEEE, 2000
15. M. Litoiu, J. Rolia, G. Serazzi, Designing process replication and activation: a quantitative approach, IEEE Transactions on Software Engineering, vol. 26, no. 12, pp. 1168-1178, 2000
16. V. Marangozova, D. Hagimont, An infrastructure for CORBA component replication, Proceedings of the 1st IFIP/ACM Working Conference on Component Deployment, Lecture Notes in Computer Science, Springer-Verlag, 2002, pp. 222-232

17. M. Marzolla, Simulation-based performance modeling of UML software architectures, Dottorato di Ricerca in Informatica, II Ciclo Nuova Serie, Dipartimento di Informatica, Università Ca' Foscari di Venezia, 2003

18. P. Narasimhan, L. E. Moser, P. M. Melliar-Smith, Enforcing determinism for the consistent replication of multithreaded CORBA applications, Proceedings of the 18th IEEE Symposium on Reliable Distributed Systems (SRDS 99), Lausanne, Switzerland, 1999, pp. 263-273

19. P. Narasimhan, L. E. Moser and P. M. Melliar-Smith, Strong replica consistency for fault tolerant CORBA applications, Journal of Computer Systems Science and Engineering, CRL Publishing, 2002

20. V. F. Nicola, P. Shahabuddin and M. Nakayama, Techniques for the fast simulation of models of highly dependable systems, IEEE Transactions on Reliability, 50, 3, 2001, pp. 246-264

21. Object Management Group, Fault tolerant CORBA, OMG Technical Committee Document, 2001-09-29, September 2001

22. Object Management Group, Object Management Architecture Guide, revision 3.0, OMG Technical Committee Document ab/97-05-05, June 1995

23. Object Management Group, The Common Object Request Broker: Architecture and Specification, revision 2.3.1, OMG Technical Committee Document formal/99-10-07, October 1999

24. B. Ramamurthy, S. J. Upadhyaya, R. K. Iyer, An object-oriented test-bed for the evaluation of checkpointing and recovery systems, Proceedings of the 27th International Symposium on Fault-Tolerant Computing, IEEE, Seattle, WA, USA, 1997, pp. 194-203

25. Y. Ren, D. E. Bakken, T. Courtney, M. Cukier, D. A. Karr, P. Rubel, C. Sabnis, W. H. Sanders, R. E. Schantz, M. Seri, AQuA: An Adaptive Architecture that Provides Dependable Distributed Objects, IEEE Transactions on Computers, Vol. 52, No. 1, 2003, pp. 31-50

26. R. D. Schlichting, F. B. Schneider, Fail-Stop processors: An approach to designing fault-tolerant computing systems, ACM Transactions on Computer Systems, 1, 3, 1983

27. D. C. Schmidt, Evaluating architectures for multithreaded object request brokers, Communications of the ACM, vol. 41, no. 10, pp. 54-60, 1998

28. T. Schnekenburger, Load balancing in CORBA: A survey of concepts, patterns and techniques, The Journal of Supercomputing, 15, 141-161, Kluwer Academic, 2000

29. P. Narasimhan, L. E. Moser, P. M. Melliar-Smith, Strongly Consistent Replication and Recovery of Fault-Tolerant CORBA Applications, Journal of Computer System Science and Engineering, Spring 2002

30. Theodoros Soldatos and Nantia Iakovidou, Performance Tradeoffs in Policies for Application Level Fault Tolerance, International Conference on Computer Sciences, Software Engineering, Information Technology, E-Business and Applications (CSITeA ’04), I.S.B.N.: 0.9742059.1.5, Cairo – Egypt, December 2004