

ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ  
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ  
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ



## ΤΥΠΙΚΗ ΑΝΑΛΥΣΗ ΚΑΙ ΑΠΟΔΟΣΗ ΑΡΧΙΤΕΚΤΟΝΙΚΩΝ ΚΑΤΑΝΕΜΗΜΕΝΩΝ ΣΥΝΑΛΛΑΓΩΝ

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ  
ΑΝΑΚΡΕΟΝΤΑΣ Π. ΜΕΝΤΗΣ

ΙΟΥΝΙΟΣ 2012



ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ  
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ  
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ



## ΤΥΠΙΚΗ ΑΝΑΛΥΣΗ ΚΑΙ ΑΠΟΔΟΣΗ ΑΡΧΙΤΕΚΤΟΝΙΚΩΝ ΚΑΤΑΝΕΜΗΜΕΝΩΝ ΣΥΝΑΛΛΑΓΩΝ

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ  
ΑΝΑΚΡΕΟΝΤΑΣ Π. ΜΕΝΤΗΣ

ΕΠΙΒΛΕΠΟΝΤΑΣ: ΠΑΝΑΓΙΩΤΗΣ ΚΑΤΣΑΡΟΣ

ΙΟΥΝΙΟΣ 2012



ΤΥΠΙΚΗ ΑΝΑΛΥΣΗ ΚΑΙ ΑΠΟΔΟΣΗ ΑΡΧΙΤΕΚΤΟΝΙΚΩΝ ΚΑΤΑΝΕΜΗΜΕΝΩΝ  
ΣΥΝΑΛΛΑΓΩΝ

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

ΑΝΑΚΡΕΟΝΤΑΣ Π. ΜΕΝΤΗΣ

Υποβλήθηκε στο Τμήμα Πληροφορικής του Αριστοτελείου Πανεπιστημίου  
Θεσσαλονίκης

ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ

Παναγιώτης Κατσαρός	Επίκουρος Καθηγητής	Α.Π.Θ.
Λευτέρης Αγγελής	Αναπληρωτής Καθηγητής	Α.Π.Θ.
Ιωάννης Σταμέλος	Αναπληρωτής Καθηγητής	Α.Π.Θ.
Αλέξανδρος Χατζηγεωργίου	Επίκουρος Καθηγητής	Πανεπιστήμιο Μακεδονίας
Απόστολος Παπαδόπουλος	Επίκουρος Καθηγητής	Α.Π.Θ.
Ελένη Καρατζά	Καθηγητής	Α.Π.Θ.
Χρήστος Γεωργιάδης	Επίκουρος Καθηγητής	Πανεπιστήμιο Μακεδονίας



© Ανακρέοντας Π. Μεντής 2012

© Α.Π.Θ 2012

## Τυπική Ανάλυση και Απόδοση Αρχιτεκτονικών Κατανεμημένων Συναλλαγών

Η έγκριση της παρούσας διδακτορικής διατριβής από το Τμήμα Πληροφορικής δεν υποδηλώνει αποδοχή των απόψεων του συγγραφέα (Ν. 5343/1932, άρθρο 202, παρ. 2)





# Ευχαριστίες

Η παρούσα διδακτορική διατριβή είναι το αποτέλεσμα ερευνητικής εργασίας που πραγματοποιήθηκε τα τελευταία χρόνια στο Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης, στο Τμήμα Πληροφορικής. Θα ήθελα να εκφράσω την ευγνωμοσύνη που νιώθω για εκείνους που συνέβαλαν με την καθοδήγηση που μου παρείχαν αυτά τα χρόνια, την εμπειρία τους, τις ερευνητικές ιδέες που μοιράστηκαν μαζί μου και τον κόπο που κατέβαλαν για τα αποτελέσματα που παρουσιάζονται στο συγκεκριμένο πόνημα.

Θέλω να ευχαριστήσω τον επιβλέποντα της διδακτορικής μου διατριβής, Επίκουρο Καθηγητή κ. Παναγιώτη Κατσαρό για την πολύτιμη επιστημονική καθοδήγηση και την εμπιστοσύνη και υποστήριξη που μου παρείχε σε όλη την διάρκεια της διδακτορικής έρευνας. Θέλω επίσης να ευχαριστήσω τον Αναπληρωτή Καθηγητή κ. Λευτέρη Αγγελή για την καθοριστική συμβολή σε ζητήματα στατιστικής ανάλυσης που συνιστούν μεγάλο μέρος της συνεισφοράς της διατριβής. Ευχαριστώ τον Αναπληρωτή Καθηγητή κ. Ιωάννη Σταμέλο για τις συμβουλές του σε ζητήματα που άπτονται της αρχιτεκτονικών συστημάτων λογισμικού. Σε εκείνον οφείλεται η γνωριμία και η μετέπειτα συνεργασία με τον Καθ. Εφαρμογών κ. Γιώργο Κακαρόντζα.

Τέλος, ευχαριστώ θερμά τους πολύ καλούς μου φίλους, Σταμάτη Πουλιάση και Στέφανο Δρένο, που επιμελήθηκαν το κείμενο της διατριβής και συνέβαλαν στη βελτίωση της ποιότητας του.



# Περίληψη

Τα σύγχρονα πληροφοριακά συστήματα χαρακτηρίζονται από κατανεμημένη αρχιτεκτονική που συνήθως περιλαμβάνει πολλαπλούς εξυπηρετητές εφαρμογών. Ο σχεδιασμός των εφαρμογών βασίζεται στις θεμελιώδεις αρχές της αντικειμενοστρέφειας και η αρχιτεκτονική του συστήματος ταυτίζει το λογικό επιμερισμό της εφαρμογής σε διάφορα επίπεδα (π.χ εμφάνιση, λογική επεξεργασία, δεδομένα) με τη δομή του λογισμικού και του υλικού που το συναποτελούν. Οι συναλλαγές παρέχουν ένα κατάλληλο μηχανισμό για την οργάνωση κατανεμημένων εργασιών λόγω των ιδιοτήτων ACID (Ατομικότητα, Συνέπεια, Απομόνωση, Μονιμότητα) που παρέχουν. Ωστόσο, οι περιορισμοί που επιβάλλουν οι ACID ιδιότητες είναι πολύ δεσμευτικοί για κάποιες εφαρμογές. Σε περιβάλλον κινητού υπολογισμού ή σε εφαρμογές ιστού, λόγου χάρη, παρατηρούνται σχετικά μεγάλες καθυστερήσεις στην επικοινωνία ή συχνές απώλειες σύνδεσης που παρατείνουν το χρόνο εκτέλεσης των συναλλαγών και προσωρινά αποκλείουν άλλες ταυτόχρονες συναλλαγές από την πρόσβαση σε διαμοιρασμένους πόρους. Στα σύγχρονα μοντέλα συναλλαγών κάποιες από τις ACID ιδιότητες αντιμετωπίζονται λιγότερο αυστηρά ώστε να απαντηθούν οι προκλήσεις που παρουσιάζουν τα προαναφερθέντα περιβάλλοντα εκτέλεσης. Στη μνήμη συναλλαγών, λόγου χάρη, δεν παρέχονται οι εγγυήσεις συνέπειας και μονιμότητας. Τα νέα μοντέλα συναλλαγών, που συχνά αποκαλούνται προηγμένες συναλλαγές, ενδέχεται να μην ικανοποιούν κάποιες από τις επιδιωκόμενες ιδιότητες ορθότητας. Ο τυπικός ορισμός και τυπικός έλεγχος των ιδιοτήτων είναι θεμελιώδεις για την εγκυρότητα ενός μοντέλου συναλλαγών, αν και σήμερα η εγκυρότητα διασφαλίζεται κυρίως με προσεγγίσεις θεωρητικής απόδειξης παρά τις πρό-

σφατες εξελίξεις στο πεδίο του έλεγχου μοντέλων. Επιπλέον, μέχρι στιγμής δεν έχουμε δει τεχνικές αυτόματης σύνθεσης αποδεδειγμένα ορθής υλοποίησης από ελεγμένο μοντέλο συναλλαγών. Η απόδοση και η διαθεσιμότητα της επεξεργασίας συναλλαγών εξαρτώνται από την αρχιτεκτονική σχεδίαση που υιοθετείται για ένα μοντέλο συναλλαγών. Οι σχεδιαστικές αποφάσεις έχουν μετρήσιμες επιδράσεις στην απόδοση και τη διαθεσιμότητα του συστήματος επεξεργασίας συναλλαγών. Τα σημεία αντιστάθμισης που υπάρχουν μεταξύ των σχεδιαστικών παραγόντων της αρχιτεκτονικής επηρεάζουν την ποιότητα της επεξεργασίας συναλλαγών με σύνθετο τρόπο.

Για την αντιμετώπιση των παραπάνω προβλημάτων προτείνουμε μια προσέγγιση καθοδηγούμενη από μοντέλα, για την αρχιτεκτονική συστημάτων συναλλαγών με δύο βασικές συνεισφορές.

Πρώτα, παρέχουμε μία συστηματική διαδικασία για τον έλεγχο μοντέλου συναλλαγών και την αυτόματη σύνθεση εγγυημένα σωστής υλοποίησης. Τα μοντέλα ορίζονται με αυτόματα που αντιστοιχούν στις οντότητες που αλληλεπιδρούν και συγχρονίζονται μεταξύ τους ως προς ένα σύνολο μηνυμάτων. Ελέγχονται όλα τα πιθανά μονοπάτια εκτέλεσης των συγχρονισμένων αυτομάτων, για τυχόν παραβίαση ιδιοτήτων της συναλλαγής.

Στη συνέχεια, επικεντρωνόμαστε σε μία προσέγγιση για τον εντοπισμό σημείων αντιστάθμισης στην αρχιτεκτονική και την εκτίμηση της επίδρασης των σχεδιαστικών αποφάσεων στην απόδοση και τη διαθεσιμότητα επεξεργασίας. Η προτεινόμενη μέθοδος εφαρμόζεται σε πειραματικά δεδομένα από προσομοίωση ή από πρωτότυπη υλοποίηση του συστήματος. Αν οι στόχοι απόδοσης και διαθεσιμότητας δεν ικανοποιούνται στα εξεταζόμενα σενάρια εκτέλεσης, η εφαρμογή της μεθόδου επαναλαμβάνεται μέχρι η αρχιτεκτονική να ικανοποιεί τους επιδιωκόμενους στόχους.

Οι παραπάνω συνεισφορές εντάσσονται στο πλαίσιο μιας ολοκληρωμένης διαδικασίας ανάπτυξης συστημάτων επεξεργασίας συναλλαγών που πιστεύουμε ότι έχει την προοπτική να αντιμετωπίσει την εγγενή πολυπλοκότητα που χαρακτηρίζει την ανάπτυξη των συστημάτων αυτών.

# Abstract

Modern network centric information systems implement highly distributed architectures that usually include multiple application servers. Application design is mainly based on the fundamental object-oriented principles and the adopted architecture matches the logical decomposition of applications into several tiers such as presentation, logic and data to their software and hardware structuring. Transactions offer a convenient mechanism for structuring distributed workloads due to the ACID properties (Atomicity, Consistency, Isolation, Durability). For some applications however, the ACID properties are too restrictive. In mobile computing or in web applications for example, long intervals of non-responsiveness or frequent disconnections lead to long-running transactions that may block concurrently running transactions and render acquired resources unavailable to other processes. New transaction models often relax one of the ACID properties in order to cope with the challenges of the aforementioned execution environments. In transactional memory for example, consistency and durability guarantees are not provided. The new transaction models, often termed as advanced transactions, may not fulfill some intended properties. Formal specification and reasoning for transaction properties is essential for proving the validity of transaction models, but it has been limited to proof-theoretic approaches despite the recent progress in model checking. Moreover, we have not seen so far automated techniques that can derive a correct implementation from a valid transaction model. Performance and availability of transaction processing depends on the adopted architectural design for a given transaction model. Design decisions have measurable effects on availability and perfor-

mance of transaction processing systems. Existing trade-offs among the design factors of an architecture affect the quality of the transaction processing service in a complex manner.

In order to address the discussed problems, we introduce a model-driven architecture based approach with a two-fold contribution.

First, we provide a systematic process for model checking a transaction model and automatically generate a correct implementation. Models are specified by state machines that represent interacting transactional entities synchronized on a set of events. Verification is performed by examining all possible execution paths of the synchronized state machines for property violations.

Second, we elaborate on an approach for discovering architecture trade-offs and for evaluating the impact of design decisions on performance and availability. The proposed method is applicable on experimental data obtained by simulation or from a prototype implementation. If an architecture design fails to fulfill the performance and availability goals for the considered workloads, the method is iteratively applied towards modifying the initial architecture until the goals are achieved.

The aforementioned contributions are part of a process for the development of transaction systems that we believe is capable to cope with the inherent complexity of distributed transaction processing.

# Περιεχόμενα

Ευχαριστίες . . . . .	v
Περίληψη . . . . .	vii
Abstract . . . . .	ix
<b>1 Εισαγωγή</b>	<b>1</b>
1.1 Κατανεμημένες συναλλαγές . . . . .	2
1.2 Ανοικτά προβλήματα και ερευνητικά θέματα . . . . .	4
1.3 Συνεισφορά της διατριβής . . . . .	6
1.4 Δομή της διατριβής . . . . .	8
<b>2 Τεχνολογίες αιχμής αρχιτεκτονικών επεξεργασίας συναλλαγών</b>	<b>11</b>
2.1 Ανάλυση αρχιτεκτονικών συστημάτων λογισμικού . . . . .	12
2.2 Προβλήματα σχεδίασης επεξεργασίας κατανεμημένων συναλλαγών . . . . .	16
2.3 Τυπική ανάλυση επεξεργασίας κατανεμημένων συναλλαγών . . . . .	20
<b>3 Έλεγχος μοντέλων και σύνθεση κώδικα επεξεργασίας κατανεμημένων συναλλαγών</b>	<b>23</b>
3.1 Ανάπτυξη συστημάτων επεξεργασίας συναλλαγών καθοδηγούμενη από μοντέλα πεπερασμένων καταστάσεων . . . . .	26
3.2 Το μοντέλο συναλλαγών δύο φάσεων (2PC) . . . . .	27
3.3 Μοντέλο ενθυλακομένων συναλλαγών . . . . .	30

3.4	Ορισμός ρόλων . . . . .	32
3.5	Ορισμός μηνυμάτων συναλλαγών . . . . .	35
3.6	Πλήρης διερεύνηση σεναρίων εκτέλεσης . . . . .	38
3.6.1	Αλγόριθμος path_explore . . . . .	41
3.7	Επαλήθευση ιδιοτήτων μοντέλων κατανεμημένων συναλλαγών . . . . .	44
3.7.1	Ιδιότητες μοντέλων κατανεμημένων συναλλαγών . . . . .	46
3.8	Σύνθεση κώδικα . . . . .	49
3.9	Συμπεράσματα . . . . .	51
<b>4</b>	<b>Εκτίμηση απόδοσης συστημάτων επεξεργασίας συναλλαγών μέσω προσομοίωσης</b>	<b>53</b>
4.1	Διαδικασία ανάπτυξης μοντέλων προσομοίωσης . . . . .	55
4.2	Αφαιρέσεις προσομοίωσης κατανεμημένων συναλλαγών . . . . .	58
4.3	Ο προσομοιωτής ACID Sim Tools . . . . .	61
4.4	Εκτίμηση απόδοσης μοντέλου κατανεμημένων συναλλαγών . . . . .	64
4.4.1	Βασικά ενδεχόμενα εκτέλεσης . . . . .	65
4.4.2	Εκτίμηση απόδοσης και διαθεσιμότητας . . . . .	65
4.4.3	Προσομοιωμένος συνθετικός φόρτος εργασίας . . . . .	66
4.4.4	Επικύρωση μοντέλου προσομοίωσης . . . . .	68
4.4.5	Πειραματικός σχεδιασμός . . . . .	69
4.4.6	Στατιστική ανάλυση . . . . .	70
4.5	Συμπεράσματα . . . . .	77
<b>5</b>	<b>Παράγοντες με σύνθετες αλληλεπιδράσεις στην απόδοση κατανεμημένων συναλλαγών</b>	<b>78</b>
5.1	Διαδικασία εκτίμησης αρχιτεκτονικών συναλλαγών . . . . .	80
5.2	Ανάλυση παραγόντων και λανθάνοντα χαρακτηριστικά ποιότητας . . . . .	83
5.2.1	Πειραματικός σχεδιασμός . . . . .	84



5.2.2	Ανάλυση παραγόντων . . . . .	86
5.3	Συνθετικές μετρικές απόδοσης επεξεργασίας συναλλαγών . . . . .	90
5.4	Ανάλυση διακύμανσης για τη διερεύνηση σχέσεων αντιστάθμισης στην απόδοση κατανεμημένων συναλλαγών . . . . .	94
5.4.1	Ανάλυση διακύμανσης . . . . .	96
5.4.2	Γραφική απεικόνιση σχέσεων αντιστάθμισης . . . . .	107
5.4.3	Απειλές στην εγκυρότητα των αποτελεσμάτων . . . . .	109
5.5	Συμπεράσματα . . . . .	111
<b>6</b>	<b>Επίλογος</b>	<b>113</b>
6.1	Σύνοψη της ερευνητικής συνεισφοράς . . . . .	114
6.2	Ερευνητικές προοπτικές . . . . .	116
<b>Appendix</b>		
	Παράρτημα . . . . .	120
I.1	Χρήση του εργαλείου ACID Sim Tools στα πλαίσια της διαδικασίας αποτίμησης αρχιτεκτονικής συναλλαγών . . . . .	120
<b>Βιβλιογραφία</b>		<b>122</b>

# Κατάλογος σχημάτων

3.1	Επιβεβαίωση και υλοποίηση μοντέλων συναλλαγών . . . . .	25
3.2	Απλοποιημένο διαγράμματα του αυτόματου ρόλου μετόχου στο βασικό πρωτόκολλο συναλλαγών δύο φάσεων . . . . .	28
3.3	Απλοποιημένα διαγράμματα αυτόματων των συντονιστών ενθυλακομένων συναλλαγών . . . . .	31
3.4	Τμήμα προδιαγραφής ρόλων του μοντέλου ενθυλακομένων συναλλαγών . .	34
3.5	Τμήμα προδιαγραφής μηνυμάτων συναλλαγών . . . . .	36
3.6	Περιορισμοί στην παραγωγή μηνυμάτων συναλλαγών . . . . .	37
3.7	Γράφημα προσεγγισιμότητας του συγχρονισμένου γινομένου των αυτομάτων των ρόλων . . . . .	39
3.8	Ιδιότητα ασφαλείας: Οι μέτοχοι συναλλαγών εφαρμόζουν την απόφαση του συντονιστή . . . . .	47
3.9	Ιδιότητα ασφαλείας: Οι συναλλαγές ολοκληρώνονται (οριστικοποίηση ή ακύρωση αποτελεσμάτων) . . . . .	48
3.10	Ιδιότητα ασφαλείας: Οι υποσυναλλαγές ακυρώνονται αν ακυρωθεί η συναλλαγή . . . . .	48
3.11	Ιδιότητα ασφαλείας: Η συναλλαγή μονιμοποιεί τα αποτελέσματα της αφού έχουν ολοκληρωθεί οι υποσυναλλαγές της . . . . .	48

3.12 Ιδιότητα προσεγγισιμότητας: Οι υποσυναλλαγές ακυρώνονται ανεξάρτητα από την έκβαση της συναλλαγής . . . . .	49
3.13 Ιδιότητα βιωσιμότητας: Η συναλλαγή οριστικοποιεί ή ακυρώνει τα αποτελέσματα της . . . . .	49
3.14 Πρότυπο σύνθεσης κώδικα για την εκτέλεση συναλλαγών . . . . .	50
4.1 Αποτίμηση μέσω προσομοίωσης στόχων απόδοσης και διαθεσιμότητας σε συστήματα επεξεργασίας ασύγχρονων μηνυμάτων . . . . .	56
4.2 UML διάγραμμα διεπαφής του αντικειμένου obj1 . . . . .	59
4.3 Διάγραμμα ακολουθίας για τη συναλλαγή tr1 . . . . .	60
4.4 Διάταξη των αρθρωμάτων προσομοίωσης (αριστερή πλευρά) και το διαδραστικό περιβάλλον προσομοίωσης του OMNET++ (δεξιά πλευρά) . . . . .	63
4.5 Η κατανομή της συχνότητας εμφάνισης των τιμών της διαμεταγωγής κατανεμημένων συναλλαγών πριν και μετά το μετασχηματισμό Blom . . . . .	71
4.6 Ραβδογράμματα των βασικών παραγόντων και οι επιδράσεις τους στην απόδοση και τη διαθεσιμότητα εξυπηρέτησης . . . . .	74
4.7 Επίδραση πρωτοκόλλου συναλλαγών, συχνότητας λήψης σημείων ελέγχου και αριθμού νημάτων επεξεργασίας στη διαμεταγωγή τοπικών συναλλαγών . . . . .	75
4.8 Η επίδραση στη διαθεσιμότητα του εξυπηρετητή acp2 του πρωτοκόλλου συναλλαγών, συχνότητας λήψης σημείων ελέγχου και αριθμού νημάτων επεξεργασίας . . . . .	76
5.1 Προτεινόμενη μέθοδος στα πλαίσια της γενικής διαδικασίας αποτίμησης αρχιτεκτονικών . . . . .	81
5.2 Ισχυρά αρνητική συσχέτιση των μεταβλητών trut_local και blocking_distr στα αρχικά και τα μετασχηματισμένα δεδομένα . . . . .	87
5.3 Επιπτώσεις των πειραματικών παραγόντων στις τρεις συνιστώσες . . . . .	95
5.4 Επιπτώσεις των πειραματικών παραγόντων στις τρεις συνιστώσες . . . . .	97

5.5	Οι αλληλεπιδράσεις μεταξύ των παραγόντων CI και MITofSF στη συνιστώσα C1 . . . . .	100
5.6	Οι αλληλεπιδράσεις μεταξύ των παραγόντων MPL, TT και MIT στη συνιστώσα C2 . . . . .	103
5.7	Οι αλληλεπιδράσεις μεταξύ των παραγόντων TT και MIT στη συνιστώσα C3	104
5.8	Διαγράμματα για την επίδραση του παράγοντα TT στις τρεις συνιστώσες .	108

# Κατάλογος πινάκων

2.1	Χαρακτηριστικά ποσοτικών μεθόδων ανάλυσης αρχιτεκτονικών με σημεία αντιστάθμισης . . . . .	14
3.1	Συναρτήσεις για τον ορισμό ιδιοτήτων ορθότητας . . . . .	45
4.1	Κλάσεις συναλλαγών, καλούμενες μέθοδοι και τρόπος πρόσβασης αντικειμένων . . . . .	66
4.2	Απαιτήσεις σε υπολογιστικούς πόρους και μνήμη των διαχρονικών αντικειμένων . . . . .	68
4.3	Παράμετροι των εξυπηρετητών συναλλαγών . . . . .	69
4.4	Πειραματικοί παράγοντες που επηρεάζουν την απόδοση και τη ταχύτητα ανάνηψης από αποτυχίες εξυπηρέτησης . . . . .	70
4.5	Αλληλεπιδράσεις παραγόντων χωρίς σημαντικές επιπτώσεις στην απόδοση και τη διαθεσιμότητα εξυπηρέτησης . . . . .	72
4.6	Παράγοντες με επιπτώσεις στη διαθεσιμότητα και την απόδοση συστημάτων επεξεργασίας καταναμημένων συναλλαγών . . . . .	73
5.1	Μετρικές απόδοσης για τον εξεταζόμενο συνθετικό φόρτο εργασίας . . . . .	84
5.2	Ανεστραμμένος πίνακας συνιστωσών της μεθόδου ανάλυσης παραγόντων	89
5.3	Ομάδες ποιοτικών μετρικών της ανάλυσης παραγόντων . . . . .	89
5.4	Στατιστικά σημαντικές επιδράσεις της συνιστώσας C1 . . . . .	99

5.5	Στατιστικά σημαντικές διαφορές μεταξύ των παραγόντων της συνιστώσας C1	100
5.6	Στατιστικά σημαντικές επιδράσεις της συνιστώσας C2 (μετρικές χρόνου απόκρισης) . . . . .	102
5.7	Στατιστικά σημαντικές διαφορές μεταξύ των παραγόντων της συνιστώσας C2 (μετρικές χρόνου απόκρισης) . . . . .	103
5.8	Στατιστικά σημαντικές επιδράσεις της συνιστώσας C3 . . . . .	105
5.9	Στατιστικά σημαντικές διαφορές μεταξύ των παραγόντων της συνιστώσας C3	106

Αγγλικός όρος	Ελληνική μετάφραση
checkpoint	σημείο ελέγχου
trade-of	σχέση αντιστάθμισης
commit	οριστικοποίηση αποτελεσμάτων
abort	αναίρεση μεταβολών
throughput	διαμεταγωγή
latent	λανθάνων
loadings	τιμές
reification	πράγμωση
factorial analysis of variance	παραγοντική ανάλυση διακύμανσης
liveness	βιωσιμότητα
reachability	προσεγγισιμότητα
pattern	πρότυπο
reification	απόδοση υπόστασης
fail-stop failure	αποτυχία πτώσης συστήματος
deadlock	αδιέξοδο

# Κεφάλαιο 1

## Εισαγωγή

Στη παρούσα διατριβή, εξετάζουμε με συστηματικό τρόπο τις επιδράσεις που έχουν οι αποφάσεις αρχιτεκτονικής σχεδίασης σε συστήματα επεξεργασίας κατανεμημένων συναλλαγών.

Η περιγραφή της αρχιτεκτονικής ενός συστήματος λογισμικού καταγράφει τα συστατικά μέρη που το αποτελούν, τις μεταξύ τους αλληλεπιδράσεις, καθώς και τους κανόνες σύνθεσης των συστατικών [Shaw and Garlan, 1996]. Η απόδοση του συστήματος καθορίζεται σε μεγάλο βαθμό από τις σχεδιαστικές επιλογές αλλά επηρεάζεται και από εξωτερικούς παράγοντες που εκδηλώνονται κατά τη λειτουργία του συστήματος καθώς και από επιλογές υλοποίησης όπως η γλώσσα προγραμματισμού και οι δομές δεδομένων που χρησιμοποιούνται. Παρότι η αρχιτεκτονική δε μπορεί να χρησιμοποιηθεί για την εξαγωγή προβλέψεων ακριβείας [Dobrica and Niemelä, 2002], αυτή αποτελεί βασικό πλαίσιο αναφοράς για την ανάλυση των επιπτώσεων που έχουν οι σχεδιαστικές επιλογές σε ποιοτικά χαρακτηριστικά όπως η απόδοση, η αξιοπιστία του συστήματος και η διαθεσιμότητα εξυπηρέτησης.

Οι επιλογές σχεδίασης της αρχιτεκτονικής περιορίζουν τα πλαίσια διακύμανσης των προαναφερθέντων ποιοτικών χαρακτηριστικών. Για την αντικειμενική εξέταση των επιδράσεων που έχουν στην ποιότητα εξυπηρέτησης, είναι απαραίτητο να οριστούν μετρικές που ποσοτικοποιούν συγκεκριμένες πτυχές της συμπεριφοράς κατά τη λειτουργία του συστή-



ματος.

Σε γενικές γραμμές, οι σχεδιαστές λογισμικού γνωρίζουν τους παράγοντες που επηρεάζουν την ποιότητα εξυπηρέτησης στα συστήματα που αναπτύσσουν. Ωστόσο, ο προσδιορισμός των παραγόντων που είναι καθοριστικοί για τις μετρικές ποιότητας και τις σύνθετες αλληλεπιδράσεις στο πλαίσιο της αρχιτεκτονικής του κάθε συστήματος παραμένει ένα ανοικτό ερευνητικό ζήτημα. Ενδεχόμενη αλλαγή στη τιμή ενός παράγοντα μπορεί να επηρεάσει πολλές μετρικές. Ορισμένες πτυχές της συμπεριφοράς βελτιώνονται ενώ κάποιες άλλες μπορεί να χειροτερεύσουν.

## 1.1 Κατανεμημένες συναλλαγές

Τα πρωτόκολλα κατανεμημένων συναλλαγών χρησιμοποιήθηκαν αρχικά κατά την ανάπτυξη προηγμένων συστημάτων βάσεων δεδομένων. Η επιτυχία των συναλλαγών και η ευρύτατη χρήση τους οφείλεται στις εγγυήσεις που παρέχουν για τις επιπτώσεις εκτέλεσής τους. Η ατομικότητα εκτέλεσης (atomicity) εγγυάται ότι σε περίπτωση σφαλμάτων επεξεργασίας, η συναλλαγή δεν έχει καμία επίπτωση στα δεδομένα. Η συνέπεια (consistency) διασφαλίζει ότι η συναλλαγή μεταφέρει τη βάση από μία συνεπή κατάσταση σε άλλη επίσης συνεπή κατάσταση. Δε μπορεί δηλαδή να παραβιαστεί η ακεραιότητα των δεδομένων από την εκτέλεση συναλλαγών. Η απομόνωση (isolation) εξασφαλίζει ότι οι επιπτώσεις μιας συναλλαγής δεν είναι ορατές σε άλλες συναλλαγές που εκτελούνται ταυτόχρονα. Τέλος, η μονιμότητα (durability) εγγυάται ότι τα αποτελέσματα των ολοκληρωμένων συναλλαγών παραμένουν μετά από την εκτέλεσή τους, ανεξάρτητα από την κατάσταση λειτουργίας του συστήματος.

Παραλλαγές του αρχικού πρωτοκόλλου ατομικής εκτέλεσης κατανεμημένων συναλλαγών βρίσκουν εφαρμογές σε πληθώρα συστημάτων κατανεμημένης επεξεργασίας. Σε αντικειμενοστραφή συστήματα λογισμικού, οι προδιαγραφές του OMG [OMG, 2003] επεκτείνουν τα αντικείμενα με λειτουργικότητα που υποστηρίζει την ατομικότητα εκτέλεσης των

μεθόδων των αντικειμένων και τη μονιμότητα των μεταβολών που προκαλεί η επεξεργασία συναλλαγών στην κατάσταση τους.

Η εφαρμογή των πρωτοκόλλων επεξεργασίας συναλλαγών σε περιβάλλοντα με ιδιαίτερες απαιτήσεις ανέδειξε ορισμένες αδυναμίες τους. Οι ACID ιδιότητες απαιτούν πόρους που σε ορισμένες συνθήκες εκτέλεσης είτε δε μπορούν να διατεθούν είτε το όφελος που αποδίδει μία ιδιότητα δε δικαιολογεί το κόστος που απαιτεί η υλοποίησή της. Σε διαδικτυακά συστήματα λόγου χάρη, ανάμεσα σε διαδοχικές αλληλεπιδράσεις των χρηστών με το σύστημα παρεμβάλλονται μεγάλα και κατά βάση απρόβλεπτα χρονικά διαστήματα. Σε εφαρμογές κινητής τηλεφωνίας, σημειώνονται συχνές αποσυνδέσεις των χρηστών από το δίκτυο. Σε περιβάλλοντα με απρόβλεπτη καθυστέρηση στην επικοινωνία ή με συχνές διακοπές σύνδεσης, τα αντικείμενα που δεσμεύει μία συναλλαγή δεν είναι διαθέσιμα σε άλλες διεργασίες έως ότου ο χρήστης επανασυνδεθεί στο σύστημα. Επιπλέον, οι συσκευές τηλεφωνίας έχουν αυστηρούς περιορισμούς στη διάθεση ενέργειας, σε υπολογιστική ισχύ και μνήμη με αποτέλεσμα να μην είναι σε θέση να υλοποιήσουν αποδοτικά όλες τις πτυχές των κλασικών πρωτοκόλλων επεξεργασίας συναλλαγών. Για την επίλυση των προβλημάτων επεξεργασίας συναλλαγών σε απαιτητικά περιβάλλοντα εκτέλεσης, μία ή περισσότερες από τις ACID ιδιότητες αγνοούνται ή αντιμετωπίζονται λιγότερο αυστηρά.

Σημαντική εξέλιξη είναι επίσης το πρόσφατο έντονο ερευνητικό ενδιαφέρον για τις μνήμες συναλλαγών (transactional memory) [Herlihy and Moss, 1993]. Ο παράλληλος προγραμματισμός με κλασικά μέσα συγχρονισμού όπως σηματοφόροι (semaphors), κοινή μνήμη (shared memory) και αποκλειστική χρήση μεταβλητών (locking) είναι επιρρεπής σε αδιέξοδα. Οι μνήμες συναλλαγής είναι μία εξαιρετικά χρήσιμη αφαίρεση που διευκολύνει την αποφυγή των προαναφερθέντων προβλημάτων. Αντίθετα από τα κλειδιά (locks) και τους σηματοφόρους, οι μνήμες συναλλαγών μπορούν να συντεθούν σε πολύπλοκες δομές χωρίς να απαιτείται γνώση του τρόπου λειτουργίας των επιμέρους συναλλαγών. Σημαντική διαφορά σε σχέση με τις κλασικές συναλλαγές είναι ότι η μονιμότητα των αποτελεσμάτων δε σχετίζεται με το τρόπο λειτουργίας της προσωρινής μνήμης και γι' αυτό δεν υλοποιείται. Η σύνθεση συναλλαγών [Harris et al., 2008] έχει επιτευχθεί για την ώρα μόνο σε συναρτη-

σιακές γλώσσες προγραμματισμού που έχουν σχετικά περιορισμένη διάδοση στη κοινότητα των προγραμματιστών.

## 1.2 Ανοικτά προβλήματα και ερευνητικά θέματα

Τα τροποποιημένα πρωτόκολλα συναλλαγών που έχουν προταθεί στη σχετική βιβλιογραφία αναφέρονται σε επεξεργασία προηγμένων συναλλαγών που έχουν κοινό γνώρισμα τη χαλαρή αντιμετώπιση ή κατάργηση κάποιων από τις ACID ιδιότητες. Ωστόσο, η κατάργηση ή τροποποίηση οποιασδήποτε ιδιότητας μπορεί να επηρεάσει τις εγγυήσεις που υποτίθεται ότι παρέχει ένα πρωτόκολλο επεξεργασίας συναλλαγών.

Το ACTA [Chrysanthis and Ramamritham, 1994] στηρίζεται σε μία αξιωματική θεμελίωση πρωτοβάθμιας λογικής για την περιγραφή και τη σύνθεση πρωτοκόλλων συναλλαγών καθώς και την απόδειξη ιδιοτήτων τους. Αλλάζοντας κάποιο από τα αξιώματα που ορίζουν το κλασσικό πρωτόκολλο συναλλαγών οριστικοποίησης δύο φάσεων (2PC) μπορούμε να περιγράψουμε άλλα πρωτόκολλα συναλλαγών και να μελετήσουμε συστηματικά τις επιδράσεις που έχει η συγκεκριμένη αλλαγή στην προδιαγραφή των ιδιοτήτων τους. Ωστόσο, το ACTA δε παρέχει εργαλεία που να διευκολύνουν τον ορισμό της συμπεριφοράς και την επαλήθευση των ιδιοτήτων του πρωτοκόλλου. Επιπλέον, η πρωτοβάθμια λογική δεν είναι κατάλληλη για τη διατύπωση μοντέλων εκτίμησης της απόδοσης αρχιτεκτονικών επεξεργασίας συναλλαγών.

Κανένα μέρος της διαδικασίας ACTA δεν υποστηρίζεται από αυτόματη επαλήθευση καθιστώντας τη μέθοδο αρκετά απαιτητική στην εφαρμογή της. Η επαλήθευση με έλεγχο μοντέλων (model checking) είναι μια προσέγγιση που υποστηρίζεται από πλήθος εργαλείων τα οποία αυτοματοποιούν και διευκολύνουν τη διαδικασία. Οι ιδιότητες που ελέγχονται εκφράζονται με όρους χρονικής λογικής. Ένας αλγόριθμος επαλήθευσης ελέγχει αν οι καταστάσεις που μπορεί να προσεγγίσει το μοντέλο επαληθεύουν την επιδιωκόμενη ιδιότητα και στην περίπτωση που αυτό δε συμβαίνει αναφέρει μονοπάτια εκτέλεσης που την παρα-

βιάζουν. Οι απαιτήσεις του αλγορίθμου επαλήθευσης σε μνήμη και χρόνο εκτέλεσης τον καθιστούν μη εφαρμόσιμο σε μοντέλα με πολύ μεγάλο χώρο καταστάσεων. Ωστόσο, όταν ο χώρος καταστάσεων το επιτρέπει, η επαλήθευση μοντέλων είναι πολύ ελκυστική λόγω της ευκολίας εφαρμογής της.

Με τη διασφάλιση της ορθότητας του διαφοροποιημένου πρωτοκόλλου, το σημαντικότερο ερώτημα αφορά τις επιπτώσεις που επιφέρει η αλλαγή στην απόδοση επεξεργασίας συναλλαγών. Σε επίπεδο αρχιτεκτονικής συστήματος, οι διάφορες σχεδιαστικές αποφάσεις ενδέχεται να έχουν θετικές επιπτώσεις σε ορισμένες από τις μετρικές ποιότητας που ενδιαφέρουν ενώ ταυτόχρονα να επιδεινώνουν άλλες. Ζητήματα σχεδίασης που ενέχουν αποφάσεις με ταυτόχρονες θετικές και αρνητικές επιδράσεις ονομάζονται σημεία αντιστάθμισης. Οι σύνθετες σχέσεις αλληλεπίδρασης μεταξύ των μετρικών ποιότητας σε μία αρχιτεκτονική συστήματος δε μπορεί να είναι εκ των προτέρων γνωστές αφού εξαρτώνται από την εκάστοτε αρχιτεκτονική και δε μπορούν να προβλεφθούν εμπειρικά. Επιπλέον, ενδεχόμενη αποτυχία στην εκπλήρωση συγκεκριμένων στόχων απόδοσης και διαθεσιμότητας καθιστά ιδιαίτερα ακριβή την διόρθωση της αστοχίας σε μεταγενέστερο στάδιο ανάπτυξης. Η δυνατότητα αποτίμησης της επίπτωσης των σχεδιαστικών επιλογών στην απόδοση από τα πρώτα κιόλας βήματα της ανάπτυξης λογισμικού, δηλαδή κατά το σχεδιασμό του συστήματος, αποβλέπει στη αποτροπή τέτοιων λαθών.

Η πειραματική μέτρηση της απόδοσης παρουσιάζει δυσκολίες. Καταρχάς, τα διαφοροποιημένα πρωτόκολλα πρέπει να υλοποιηθούν, και η υλοποίηση αποτελεί σημαντική επένδυση πόρων. Επιπλέον, για την εξαγωγή έγκυρων συμπερασμάτων, τα νέα πρωτόκολλα πρέπει να δοκιμαστούν σε διάφορα περιβάλλοντα εκτέλεσης. Τέλος, η επένδυση στην υλοποίηση και τη διεξαγωγή πειραμάτων θα δικαιωθεί μόνο εφόσον η νέα προσέγγιση επεξεργασίας συναλλαγών είναι αποδοτικότερη από την αρχική.

Μια πρώτη εκτίμηση της απόδοσης επεξεργασίας συναλλαγών μπορεί να αποκομισθεί μέσω προσομοίωσης. Οι οντότητες σε μοντέλα επεξεργασίας καταναμημένων συναλλαγών αλληλεπιδρούν με ασύγχρονη ανταλλαγή μηνυμάτων. Κάθε υλοποίηση των μερών ενός τέτοιου συστήματος περιλαμβάνει ένα σύνολο από μεταβλητές που καταγράφουν πληρο-

φορίες απαραίτητες για τον καθορισμό της αντίδρασης από μία οντότητα στη λήψη ενός μηνύματος. Όταν η οντότητα λάβει ένα μήνυμα αυτή μπορεί να ενημερώσει ένα υποσύνολο των μεταβλητών ή να στείλει μηνύματα σε άλλες οντότητες. Λάθη στο χειρισμό των εισερχόμενων μηνυμάτων έχουν ως αποτέλεσμα κάποιες μεταβλητές να έχουν διαφορετική τιμή από την αναμενόμενη ή προκαλούν την ετεροχρονισμένη αποστολή μηνυμάτων. Τέτοια λάθη οφείλονται στην πολυπλοκότητα του λογισμικού προσομοίωσης και ενδέχεται να αλλοιώσουν τα αποτελέσματα και να οδηγήσουν σε λάθος συμπεράσματα και αποφάσεις [Sargent, 2005]. Σε κάθε περίπτωση, το λάθος εκδηλώνεται στη συμπεριφορά του μοντέλου μετά την ανταλλαγή αρκετών μηνυμάτων όταν είναι πλέον δύσκολος ο προσδιορισμός της αιτίας που οδήγησε στην παρατηρούμενη συμπεριφορά.

### 1.3 Συνεισφορά της διατριβής

Η παρούσα διατριβή πραγματεύεται τη μελέτη της επίδρασης που έχουν οι σχεδιαστικές αποφάσεις σε επίπεδο αρχιτεκτονικής στην ποιότητα επεξεργασίας κατανεμημένων συναλλαγών. Αναπτύξαμε τεχνικές που καθοδηγούν τον αναλυτή σε όλα τα στάδια σχεδιασμού και ανάπτυξης λογισμικού επεξεργασίας συναλλαγών ώστε το τελικό προϊόν να εκπληρώνει τις προσδοκίες απόδοσης και διαθεσιμότητας εξυπηρέτησης.

Οι κυριότερες συνεισφορές της διατριβής συνοψίζονται στα παρακάτω:

1. Προτείνεται μια συστηματική διαδικασία προσομοίωσης συστημάτων στα οποία η επικοινωνία μεταξύ των μερών τους επιτυγχάνεται με ανταλλαγή ασύγχρονων μηνυμάτων [Mentis and Katsaros, 2012b]. Η πολυπλοκότητα που είναι εγγενής στα ασύγχρονα συστήματα μπορεί εύκολα να οδηγήσει σε λάθη κατά την υλοποίηση ή στην επιλογή μη βέλτιστων αρχιτεκτονικών. Η προτεινόμενη μέθοδος αξιοποιεί την επαλήθευση με έλεγχο μοντέλων για την ορθή υλοποίηση και την προσομοίωση διακριτών γεγονότων για την εκτίμηση της απόδοσης, της διαθεσιμότητας και άλλων χαρακτηριστικών ποιότητας εξυπηρέτησης.

2. Παρουσιάζουμε τον προσομοιωτή ACID Sim Tools και συνοδευτικές βιβλιοθήκες που βοηθούν στην ανάπτυξη προσομοιωτών για πρωτόκολλα κατανομημένων συναλλαγών [Mentis et al., 2008]. Τα αποτελέσματα της προσομοίωσης αξιοποιούνται για τη συγκριτική μελέτη της απόδοσης και της διαθεσιμότητας εξυπηρέτησης που αφορά το αρχικό πρωτόκολλο συναλλαγών και δύο παραλλαγών του.
3. Υλοποιήθηκαν εργαλεία που βοηθούν το σχεδιαστή αρχιτεκτονικών επεξεργασίας συναλλαγών στην επαλήθευση ιδιοτήτων ορθότητας και στην υλοποίηση μέσω αυτόματης σύνθεσης κώδικα από έγκυρο μοντέλο. Εξετάζουμε την αποτελεσματικότητα των προτεινόμενων εργαλείων στην ανάπτυξη του πρωταρχικού μοντέλου συναλλαγών (2PC) και δύο παραλλαγών του. Τα αποτελέσματα παρουσιάζονται στην εργασία [Mentis and Katsaros, 2009]. Τα μοντέλα συναλλαγών διαχωρίζονται σε επιμέρους οντότητες που αλληλεπιδρούν μέσω μηνυμάτων. Η συμπεριφορά των οντοτήτων ορίζεται με αυτόματα ενώ από τη σύνθεση των αυτομάτων προκύπτει η συμπεριφορά του μοντέλου συναλλαγών. Οι ιδιότητες που αναμένεται να εκπληρώνει το μοντέλο εκφράζονται ως περιορισμοί στα πιθανά σενάρια εκτέλεσης του σύνθετου αυτομάτου. Οι περιορισμοί είναι λογικές προτάσεις που περιέχουν λογικούς και χρονικούς τελεστές, καθώς και αναφορές σε τύπους μηνυμάτων και καταστάσεων του σύνθετου αυτομάτου. Μοντέλα συναλλαγών που επαληθεύουν τις απαιτούμενες συνθήκες ορθότητας χρησιμοποιούνται στη συνέχεια για την αυτόματη σύνθεση μέρους της υλοποίησης.
4. Ορίζονται προηγμένα μοντέλα συναλλαγών και επαληθεύονται οι ιδιότητες τους με έλεγχο μοντέλων [Mentis and Katsaros, 2012a]. Ο έλεγχος του μοντέλου και η αυτόματη σύνθεση κώδικα εγγυάται την ορθή υλοποίηση αρχιτεκτονικών επεξεργασίας προηγμένων συναλλαγών.
5. Εξετάζονται οι παράγοντες αρχιτεκτονικής σχεδίασης που επηρεάζουν χαρακτηριστικά ποιότητας στην επεξεργασία συναλλαγών, όπως η απόδοση και η διαθεσιμότητα, καθώς και οι αλληλεπιδράσεις μεταξύ των παραγόντων [Mentis et al., 2010].

Χρησιμοποιούμε στατιστικές μεθόδους για τον εντοπισμό των παραγόντων εκείνων που αλληλεπιδρούν και την ενδεχόμενη βελτίωση ή υπονόμηση που επιφέρει η συνδυασμένη επίδρασή τους σε κάποιο ποιοτικό χαρακτηριστικό. Η στατιστική ανάλυση εφαρμόζεται σε δεδομένα που λήφθηκαν με προσομοίωση ενός συνθετικού σεναρίου επεξεργασίας συναλλαγών.

6. Ορίζονται μετρικές που αντανακλούν τη συνδυασμένη επίδραση παραγόντων αρχιτεκτονικής [Mentis et al., 2009] στην ποιότητα της επεξεργασίας συναλλαγών. Προτείνονται στατιστικές μέθοδοι για την ομαδοποίηση των παραγόντων με ισχυρή επίδραση και συμβατό χαρακτήρα. Η ομαδοποίηση επιτυγχάνεται με τον εντοπισμό σχέσεων και αλληλεπιδράσεων ανάμεσα στους παράγοντες μιας ομάδας. Οι ομάδες που προκύπτουν από τη μέθοδο χρησιμοποιούνται για τον ορισμό σύνθετων μετρικών που αντανακλούν τη βελτίωση ή υπονόμηση που επιφέρουν πιθανές αλλαγές στην αρχιτεκτονική σχεδίαση του συστήματος πάνω σε ομάδες χαρακτηριστικών ποιότητας που ανιχνεύθηκαν.

Οι παραπάνω συνεισφορές εντάσσονται στο πλαίσιο μιας ολοκληρωμένης διαδικασίας ανάπτυξης συστημάτων επεξεργασίας συναλλαγών καθοδηγούμενης από μοντέλα καταστάσεων – μεταβάσεων και προσομοίωση. Τα μοντέλα καταστάσεων εξασφαλίζουν την ορθή σχεδίαση του συστήματος ως προς τις προδιαγραφές του, ενώ με την προσομοίωση διερευνάται η επίδραση των διαφόρων σχεδιαστικών παραγόντων στην απόδοση και τη διαθεσιμότητα συναλλαγών. Η διαδικασία ολοκληρώνεται με την αυτόματη σύνθεση κώδικα επεξεργασίας συναλλαγών από έγκυρα μοντέλα.

## 1.4 Δομή της διατριβής

Στο κεφάλαιο αυτό αναφερθήκαμε σε ανοικτά προβλήματα, τους ερευνητικούς στόχους και τη συμβολή της παρούσας διδακτορικής διατριβής.

Στο δεύτερο κεφάλαιο συνοψίζουμε δημοσιευμένες εργασίες που εντάσσονται στις τεχνολογίες αιχμής για τη μελέτη της αρχιτεκτονικής συστημάτων και την απόδοση επεξεργασίας κατανεμημένων συναλλαγών. Επικεντρωνόμαστε σε προβλήματα σχεδίασης μοντέλων συναλλαγών και σε πρόσφατες προσπάθειες εξέλιξής τους. Το κεφάλαιο ολοκληρώνεται με την επισκόπηση προσεγγίσεων επαλήθευσης συστημάτων επεξεργασίας συναλλαγών.

Στο τρίτο κεφάλαιο παρουσιάζουμε το βασικό πρωτόκολλο συναλλαγών 2PC, δύο παραλλαγές του και παραδείγματα προηγμένων πρωτοκόλλων συναλλαγών. Περιγράφουμε τις βασικές πτυχές της προτεινόμενης μεθόδου ελέγχου μοντέλων και την εφαρμογή της στα προαναφερθέντα πρωτόκολλα συναλλαγών. Επιδεικνύουμε τον ορισμό μοντέλων συναλλαγών με τη βοήθεια αυτομάτων, την προδιαγραφή ιδιοτήτων που πρέπει να εκπληρώνει το μοντέλο και τον αλγόριθμο που ελέγχει την εκπλήρωση αυτών. Από το μοντέλο παράγεται αυτόματα μέρος της υλοποίησης που αφορά το χειρισμό μηνυμάτων. Ο χειρισμός μηνυμάτων από τις οντότητες του συστήματος είναι ένα από τα πιο πολύπλοκα μέρη του συστήματος και συνοδεύεται από υψηλό κίνδυνο λαθών.

Στο τέταρτο κεφάλαιο παρουσιάζουμε μια συστηματική μέθοδο για την προσομοίωση μοντέλων επεξεργασίας κατανεμημένων συναλλαγών. Η μέθοδος στηρίζεται στην επαλήθευση των προδιαγραφών του μοντέλου συναλλαγών μέσω ελέγχου. Στη συνέχεια εισάγονται κριτήρια επιλογής σεναρίων εκτέλεσης που μας επιτρέπουν να διαπιστώσουμε εάν οι στόχοι απόδοσης και διαθεσιμότητας εκπληρώνονται από το υπό μελέτη μοντέλο. Οι υποθέσεις σχετικά με τις παραμέτρους προσομοίωσης αντανakλούν τις σχεδιαστικές επιλογές της αρχιτεκτονικής του συστήματος και επικυρώνονται μέσω δοκιμασμένων προσεγγίσεων. Τα δεδομένα που προκύπτουν από τις εκτελέσεις των προσομοιώσεων μπορούν να στηρίξουν μια ανάλυση ευαισθησίας για την εξαγωγή συμπερασμάτων σχετικά με την απόδοση και την διαθεσιμότητα επεξεργασίας της υπό μελέτη αρχιτεκτονικής.

Στο πέμπτο κεφάλαιο περιγράφεται μία προσέγγιση ποσοτικοποίησης των επιδράσεων που έχουν οι διάφορες σχεδιαστικές επιλογές στην ποιότητα της επεξεργασίας. Εξετάζουμε την εφαρμογή της μεθόδου μελετώντας διάφορα πρωτόκολλα συναλλαγών Επιπλέον, διε-



ρευνούμε σχέσεις αντιστάθμισης μεταξύ επιλογών σχεδίασης με στόχο την υιοθέτηση σχεδιαστικών αποφάσεων σε χαρακτηριστικά της αρχιτεκτονικής που βελτιώνουν την ποιότητα της επεξεργασίας συναλλαγών και ταυτόχρονα περιορίζουν τους κινδύνους που ελλοχεύουν στην επεξεργασία μεγάλου φόρτου συναλλαγών.

Τέλος, το έκτο κεφάλαιο συνοψίζει τα συμπεράσματα της διατριβής και τα ζητήματα που παραμένουν ανοιχτά για μελλοντική έρευνα.

## Κεφάλαιο 2

# Τεχνολογίες αιχμής αρχιτεκτονικών επεξεργασίας συναλλαγών

Οι κατακευκμμένες συναλλαγές αναπτύχθηκαν ιδιαίτερα στα πλαίσια της έρευνας στο χώρο των βάσεων δεδομένων, για την επίλυση προβλημάτων ταυτόχρονης προσπέλασης και μεταβολής δεδομένων. Παρόμοια προβλήματα υπάρχουν και σε άλλα πεδία εφαρμογών όπου ειδικά προσαρμοσμένες παραλλαγές των πρωτοκόλλων συναλλαγών διαδραματίζουν κρίσιμο ρόλο στην παροχή αξιόπιστης εξυπηρέτησης.

Τα πρωτόκολλα συναλλαγών όπως το 2PC, είναι αποδοτικά και εκτός του πλαισίου των βάσεων δεδομένων, όταν ισχύουν οι αναγκαίες συνθήκες που περιγράφονται στην εργασία [Thanisch, 2000]. Σε αυτή την εργασία εξετάζονται εφαρμογές του μοντέλου κατακευκμένων συναλλαγών σε κατακευκμένα συστήματα αρχείων, συστήματα πρόσβασης σε κατακευκμένη διαμοιρασμένη μνήμη, διαδικτυακές συναλλαγές με ανταγωνισμούς μεταξύ των μετόχων της συναλλαγής, κ.α. Η χρήση κατακευκμένων συναλλαγών είναι επωφελής όταν ισχύουν οι παρακάτω συνθήκες:

1. Απαιτείται η συμμετοχή ανεξάρτητων οντοτήτων για την ολοκλήρωση της επεξεργασίας.

2. Η επικοινωνία επιτυγχάνεται με την ανταλλαγή ασύγχρονων μηνυμάτων.
3. Οι μέτοχοι έχουν πρόσβαση σε μέρος μόνο της συνολικής κατάστασης του συστήματος.
4. Ορισμένες αποφάσεις απαιτούν γνώση της συνολικής κατάστασης και συντονισμό των δραστηριοτήτων των ανεξάρτητων οντοτήτων.

Λόγω της ευρείας διάδοσης των συναλλαγών σε μεγάλο φάσμα εφαρμογών, έχουν καταγραφεί ενδιαφέρουσες σχετικές εργασίες που εστιάζουν στην ορθή υλοποίηση πρωτοκόλλων συναλλαγών και την μελέτη της απόδοσης τους. Στις εργασίες που παρουσιάζουμε στις επόμενες ενότητες ερευνώνται τα πρωτόκολλα συναλλαγών, αλλά και υποσυστήματα απαραίτητα για την επεξεργασία κατανεμημένων συναλλαγών όπως η πρόσβαση σε κοινοποιημένους πόρους και η επαναφορά της κανονικής λειτουργίας (ανάληψη) μετά από αποτυχία πτώσης συστήματος. Γίνεται επίσης αναφορά σε σχετικές προσεγγίσεις ανάλυσης της αρχιτεκτονικής συστημάτων.

## 2.1 Ανάλυση αρχιτεκτονικών συστημάτων λογισμικού

Σε έργα πληροφορικής, η αρχιτεκτονική λογισμικού προσδιορίζεται στα αρχικά στάδια του σχεδιασμού [Grunske, 2007, Losavio, 2002, Shaw and Garlan, 1996] όταν το κόστος σχεδιαστικών αλλαγών και επιδιόρθωσης λαθών είναι πολύ μικρότερο από αντίστοιχες μεταβολές σε μεταγενέστερα στάδια της ανάπτυξης λογισμικού [Williams and Carver, 2010]. Μία διαδικασία επικύρωσης αρχιτεκτονικής [Gorton, 2006] αποτιμά τις σχεδιαστικές επιλογές αναφορικά με τα επιθυμητά ποιοτικά χαρακτηριστικά που ενδέχεται να είναι ιεραρχημένα και αλληλοεξαρτώμενα. Η εκπλήρωση ενός ποιοτικού χαρακτηριστικού ενδέχεται να υποβοηθά ή να υπονομεύει την εκπλήρωση άλλων ποιοτικών χαρακτηριστικών λόγω των

μεταξύ τους συσχετίσεων. Σχεδιαστικές επιλογές που δεν επιτρέπουν την εκπλήρωση κάποιου ποιοτικού χαρακτηριστικού θεωρούνται ριψοκίνδυνες ενώ εκείνες που συμβάλλουν στην εκπλήρωση των ποιοτικών στόχων, χωρίς να παραβιάζουν άλλες απαιτήσεις χαρακτηρίζονται ως καλές επιλογές [Grunske, 2006]. Οι σχεδιαστικές αποφάσεις θεωρούνται ριψοκίνδυνες ή ωφέλιμες για την αρχιτεκτονική ανάλογά με τις υποθέσεις για την αναμενόμενη επίδραση στη συμπεριφορά του συστήματος. Στο αρχιτεκτονικό πρότυπο του Συντονιστή Διεργασιών (Process Coordinator), λόγου χάρη, ο σχεδιαστής υποθέτει πως η απόδοση χειροτερεύει όταν οι διαδικασίες κατανέμονται σε περισσότερους υπολογιστικούς κόμβους.

Η σχετική βιβλιογραφία περιέχει εργασίες που επιχειρούν την κωδικοποίηση σε πίνακες των χαρακτηριστικών ποιότητας [Berander et al., 2006, Eguiluz and Barbacci, 2003] και των αλληλεπιδράσεων τους. Οι πίνακες προκύπτουν διαισθητικά από την εμπειρία των συγγραφέων με την αρχιτεκτονική που μελετούν και όχι από δεδομένα που καταδεικνύουν συγκεκριμένες αλληλεπιδράσεις. Σύμφωνα με την εργασία [Reussner and Firus, 2005], οι αλληλεπιδράσεις μεταξύ των ποιοτικών χαρακτηριστικών αποτελούν ιδιότητα της αρχιτεκτονικής. Η απόδοση και η αξιοπιστία λόγου χάρη, μπορεί να μεταβάλλονται αντίστροφα σε μία αρχιτεκτονική ενώ σε μία άλλη να είναι εφικτή η ταυτόχρονη βελτίωσή τους.

Ποσοτικές μέθοδοι εκτίμησης πρέπει να λαμβάνουν υπόψη ότι οι αλληλεπιδράσεις μεταξύ των ποιοτικών χαρακτηριστικών είναι συγκεκριμένες σε κάθε αρχιτεκτονική και πρέπει να παρέχουν δεδομένα που να κάνουν εμφανείς τις υπάρχουσες σχέσεις αντιστάθμισης. Πρέπει να είναι αρκετά γενικές ώστε να είναι χρήσιμες για εφαρμογή σε διάφορες αρχιτεκτονικές χωρίς να περιορίζονται μόνο σε κάποια συγκεκριμένα ποιοτικά χαρακτηριστικά και μετρικές που είναι συνιφασμένα με την εφαρμογή τους. Τέλος, είναι επιθυμητό να μπορεί να προσδιοριστεί η σχεδιαστική επιλογή που προκαλεί την παρατηρούμενη διακύμανση στις μετρήσεις κάποιου ποιοτικού χαρακτηριστικού. Αυτό επιτρέπει τον εντοπισμό σχεδιαστικών επιλογών που στη μέθοδο ATAM [Clements et al., 2001] ονομάζονται σημεία ευαισθησίας ή σημεία αντιστάθμισης. Σημείο ευαισθησίας ονομάζεται η σχεδιαστική επιλογή που επηρεάζει καθοριστικά την εκπλήρωση κάποιας απαίτησης ποιότητας. Τέτοιες επιλογές μπορεί να είναι ιδιότητα συστατικού στοιχείου της εφαρμογής ή κάποια σχέση μεταξύ συστατικών του

συστήματος [Crnkovic et al., 2004]. Το σημείο αντιστάθμισης επηρεάζει ταυτόχρονα πολλά χαρακτηριστικά ποιότητας και άπτεται σχεδιαστικών αποφάσεων ιδιαίτερα σημαντικών για την αρχιτεκτονική του συστήματος.

Εργασία	[Litoiu et al., 2000]	[Paul et al., 2003]	[Katsaros et al., 2007]
Πεδίο εφαρμογής της ποσοτικής μεθόδου	Κατανεμημένες εφαρμογές	Πρωτόκολλα σημείων ελέγχου και ανάνηψης από αποτυχία του συστήματος	Κατανεμημένα συστήματα με ανεξάρτητες διαδικασίες δημιουργίας σημείου ελέγχου
Σημείο αντιστάθμισης	Πλεόνασμα διεργασιών ή πλήθος ενεργών νημάτων επεξεργασίας	Περίοδος λήψης σημείου ελέγχου	Περίοδος λήψης σημείου ελέγχου για κάθε ανεξάρτητη δραστηριότητα
Μετρική ποιότητας	Βαθμός αξιοποίησης διεργασιών και συσκευών	Κόστος λήψης σημείων ελέγχου και ανάκαμψης	Χρόνος απόκρισης των υπολογισμών με και χωρίς σφάλματα επεξεργασίας
Στόχος της ποιοτικής μετρικής	Προσδιορισμός του επιπέδου πλεονασμού της διεργασίας ή του αριθμού ενεργών νημάτων επεξεργασίας για την αποφυγή καθυστερήσεων εξυπηρέτησης και υψηλής κατανάλωσης μνήμης	Η αποτίμηση της απόδοσης του πρωτοκόλλου σε διαφορετικά περιβάλλοντα εκτέλεσης	Ο προσδιορισμός της περιόδου λήψης σημείου ελέγχου ώστε να ικανοποιεί τους στόχους του χρόνου απόκρισης με το ελάχιστο δυνατό κόστος ανάνηψης από αποτυχίες
Μεθοδολογική προσέγγιση	Υβριδική μέθοδος που συνδυάζει αναλυτική αποτίμηση και μαθηματικό προγραμματισμό	Μετρήσεις βασισμένες σε προσομοίωση	Στατιστική ανάλυση αποτελεσμάτων προσομοίωσης
Επιδράσεις σχεδιαστικών παραγόντων σε ποιοτικές μετρικές	Μελετά μόνο τις σχεδιαστικές αποφάσεις που έχουν ληφθεί υπόψη στην κατασκευή του αναλυτικού μοντέλου. Οι εκτιμήσεις της μεθόδου αφορούν σημεία ευαισθησίας	Μελετά τις επιδράσεις όλων των σχεδιαστικών παραγόντων της αρχιτεκτονικής	Μελετά τις επιδράσεις όλων των σχεδιαστικών παραγόντων της αρχιτεκτονικής. Οι εκτιμήσεις αφορούν σημεία αντιστάθμισης

Πίνακας 2.1: Χαρακτηριστικά ποσοτικών μεθόδων ανάλυσης αρχιτεκτονικών με σημεία αντιστάθμισης

Στο Πίνακα 2.1 παρουσιάζουμε περιληπτικά τα βασικά χαρακτηριστικά κάποιων ποσοτικών μεθόδων ανάλυσης σημείων αντιστάθμισης. Αναφέρουμε τρεις διαφορετικές προσέγγι-

γησης: η πρώτη συνδυάζει αναλυτικές τεχνικές με μαθηματικό προγραμματισμό, η δεύτερη βασίζεται σε προσομοίωση και η τρίτη αξιοποιεί τη στατιστική ανάλυση αποτελεσμάτων προσομοίωσης. Οι τρεις αυτές προσεγγίσεις επικεντρώνονται σε διαφορετικά προβλήματα αντιστάθμισης. Προκύπτουν ορισμένα συμπεράσματα που τελικά οδηγούν στην ανάγκη που επιχειρεί να ικανοποιήσει η συμβολή της παρούσας διατριβής:

- Οι υπάρχουσες μέθοδοι αποτίμησης ποσοτικοποιούν αποκλειστικά τις μετρικές ποιότητας που συμπεριλαμβάνονται στην ανάλυση σημείων αντιστάθμισης. Αυτός ο περιορισμός αποκλείει την ανίχνευση επιπλέον επιδράσεων που δεν έχουν προβλεφθεί και δε συγκαταλέγονται μεταξύ των μετρικών και των σχεδιαστικών παραγόντων στους οποίους στηρίζεται η ανάλυση.
- Υποθέτουν εκ των προτέρων γνώση των σημείων αντιστάθμισης σε μία αρχιτεκτονική. Γι' αυτό το λόγο αδυνατούν να ανακαλύψουν επιπλέον σημεία ευαισθησίας ή αντιστάθμισης.
- Εκτός από τα εκ των προτέρων γνωστά σημεία αντιστάθμισης, οι προαναφερόμενες μέθοδοι αδυνατούν να αποδώσουν την παρατηρούμενη διακύμανση στις τιμές των ποιοτικών μετρικών σε συγκεκριμένες αρχιτεκτονικές επιλογές.

Η μέθοδος που εισάγουμε στο Κεφάλαιο 5, δε διέπεται από τις παραπάνω αδυναμίες στην ανάλυση σημείων αντιστάθμισης και επιπλέον παρέχει ενδείξεις και δεδομένα που αποκαλύπτουν τα άγνωστα σημεία αντιστάθμισης της αρχιτεκτονικής. Η μέθοδος υποστηρίζει την ανάλυση των αρχιτεκτονικών επιλογών από τα πρώτα κιόλας στάδια της σχεδίασης. Μετέπειτα, η γνώση που αποκομίζεται από την εφαρμογή της μεθόδου μπορεί να αξιοποιηθεί για την προσαρμογή των παραμέτρων του συστήματος σε διαμορφώσεις που εξυπηρετούν αποτελεσματικά τους στόχους ποιότητας εξυπηρέτησης.

Υποθέτουμε ότι υπάρχει η δυνατότητα πειραματισμού που παρέχει προσομοιωτής ή κάποιο benchmark [Transaction Processing Performance Council, 2005, Zhu et al., 2007] που υλοποιεί αντιπροσωπευτικό υποσύνολο των δυνατοτήτων του συστήματος. Η στατι-

στική ανάλυση στην οποία στηρίζεται η μέθοδος είναι γενική, αφού δεν περιορίζεται σε κάποιο συγκεκριμένο τύπο προσομοίωσης ή από κάποια συγκεκριμένη θεωρία για την αποτίμηση μιας μετρικής [Balsamo et al., 2004a].

## 2.2 Προβλήματα σχεδίασης επεξεργασίας κατανεμημένων συναλλαγών

Τα συστήματα επεξεργασίας κατανεμημένων συναλλαγών αποτελούνται από το πρωτόκολλο συντονισμού οριστικοποίησης των αποτελεσμάτων της συναλλαγής, τη διαχείριση της πρόσβασης σε διαμοιρασμένους πόρους, καθώς και στρατηγικές ανάκαμψης ακυρωμένων συναλλαγών και την αποκατάσταση των εξυπηρετητών μετά την εκδήλωση αποτυχιών πτώσης συστήματος (ανάληψη). Μία αποτελεσματική αρχιτεκτονική επεξεργασίας κατανεμημένων συναλλαγών οφείλει να αντιμετωπίσει τις προκλήσεις που άπτονται της σχεδίασης των επιμέρους υποσυστημάτων, έτσι ώστε το σύστημα να αποδίδει ικανοποιητικά.

Στην [Agrawal et al., 1996] εξετάζονται τρία εναλλακτικά πρωτόκολλα διαχείρισης πρόσβασης σε διαμοιρασμένους πόρους:

1. Κλείδωμα των πόρων προσωρινά για αποκλειστική χρήση. Αν εντοπιστεί αδιέξοδος μεταξύ δύο συναλλαγών που ανταγωνίζονται για τους ίδιους πόρους, εκείνη που έφτασε τελευταία, ακυρώνεται. Αναστέλλεται η εκτέλεση των συναλλαγών που χαρακτηρίζονται από συγκρούσεις για πρόσβαση σε διαμοιρασμένους πόρους μέχρι οι πόροι αυτοί να καταστούν διαθέσιμοι.
2. Κατά το δεύτερο πρωτόκολλο ακυρώνονται οι συναλλαγές που δε μπορούν να αποκτήσουν πρόσβαση σε αναγκαίους πόρους, ανεξάρτητα από το αν υπάρχουν αδιέξοδα. Οι ακυρωμένες συναλλαγές υποβάλλονται και πάλι για επεξεργασία αργότερα, όταν οι ζητούμενοι πόροι μπορεί να είναι διαθέσιμοι.
3. Η αισιόδοξη προσέγγιση στη διαχείρισης διαμοιρασμένων πόρων αποφεύγει το κό-

στος του έλεγχου για συγκρούσεις και αδιέξοδα κατά την εκτέλεση της συναλλαγής. Πριν από την οριστικοποίηση των αποτελεσμάτων, ελέγχεται αν τα δεδομένα που χρησιμοποιήσε η συναλλαγή μεταβλήθηκαν κατά την διάρκεια εκτέλεσής της από άλλη συναλλαγή. Σε αυτή την περίπτωση, η συναλλαγή ακυρώνεται και υποβάλλεται ξανά για επεξεργασία.

Εξετάζεται η απόδοση των διαφορετικών στρατηγικών διαχείρισης με βάση κάποιες μετρικές όπως η διαμεταγωγή συναλλαγών, ο χρόνος απόκρισης και το ποσοστό των συναλλαγών που ακυρώνονται λόγω συγκρούσεων για την πρόσβαση σε δεδομένα. Η μελέτη στηρίζεται σε ρητές υποθέσεις για το μοντέλο και το περιβάλλον εκτέλεσης των συναλλαγών. Αντιφατικά αποτελέσματα σε σχέση με άλλες σχετικές εργασίες αποδίδονται σε λάθος υποθέσεις για το περιβάλλον εκτέλεσης και τη συμπεριφορά των συναλλαγών. Τα πειράματα προσομοίωσης διεξάγονται για φόρτους εργασίας με διαφορετικά χαρακτηριστικά, ώστε να μελετηθεί η συμπεριφορά των πρωτοκόλλων σε διάφορες συνθήκες.

Τα πρωτόκολλα συναλλαγών έχουν σχεδιαστεί, ώστε να μπορούν να ανταπεξέρχονται σε αποτυχίες πτώσης του εξυπηρετητή συναλλαγών. Στην [Chrysanthis et al., 1998] εξετάζονται το βασικό πρωτόκολλο κατανεμημένων συναλλαγών 2PC και παραλλαγές του, ως προς την ταχύτητα ανάνηψης από αποτυχίες πτώσης. Παρουσιάζεται το μοντέλο των πρωτοκόλλων και οι υποθέσεις του και μελετάται η απόδοση με πειράματα προσομοίωσης. Εξετάζονται φόρτοι εργασίας με διαφορετικά χαρακτηριστικά όπως π.χ μόνο συναλλαγές που δε μεταβάλλουν τα δεδομένα, ώστε να μετρηθεί η βελτίωση στη διαμεταγωγή συναλλαγών, που αποφέρουν οι αλλαγές στο βασικό πρωτόκολλο.

Στην [Martin and Ramamritham, 1997] περιγράφονται φορμαλιστικά οι μηχανισμοί ανάνηψης από αποτυχίες πτώσης συστήματος. Οι ιδιότητες αποδεικνύονται με βάση το ιστορικό εκτέλεσης, δηλαδή τις πιθανές εκτελέσεις των συναλλαγών. Η εργασία είναι συμπληρωματική προς τη [Chrysanthis and Ramamritham, 1990], αλλά επικεντρώνεται στους μηχανισμούς αποκατάστασης της λειτουργίας του εξυπηρετητή και όχι στο πρωτόκολλο συναλλαγών.



Προηγμένες εκδοχές του βασικού πρωτοκόλλου συναλλαγών και εκδοχές του που είναι προσαρμοσμένα σε ειδικά περιβάλλοντα εκτέλεσης, χαλαρώνουν κάποιες από τις εγγυήσεις των συναλλαγών. Τα προηγμένα πρωτόκολλα συναλλαγών παρέχουν λύσεις πέρα από τους περιορισμούς που δίνουν τα απλά μοντέλα συναλλαγών [Elmagarmid, 1992] όπως:

1. Ότι οι συναλλαγές έχουν μικρή διάρκεια.
2. Οι συναλλαγές δεν έχουν δομή (ενθυλακωμένες συναλλαγές).
3. Δεν επιτρέπεται μερική επιτυχία της συναλλαγής (είτε οριστικοποιούνται όλα τα αποτελέσματα ή αναιρούνται όλα).
4. Οι συναλλαγές εκτελούνται απομονωμένα χωρίς δυνατότητα συνεργασίας.
5. Ο χρήστης δε μπορεί να επηρεάσει την εκτέλεσή της συναλλαγής.

Σε περιπτώσεις όπου η γρήγορη ανάνηψη από αποτυχίες πτώσης συστήματος δεν είναι εφικτή ή επιφέρει μεγάλη επιβάρυνση στην κανονική εκτέλεση των συναλλαγών, μπορούμε να αξιοποιήσουμε πλεονάζοντες εξυπηρετητές που αναλαμβάνουν την επεξεργασία συναλλαγών, όταν κάποιος από τους συμμετέχοντες εξυπηρετητές αποτύχει. Για να είναι εφικτή η αντικατάσταση του εξυπηρετητή από άλλον, πρέπει οι εσωτερικές τους καταστάσεις να είναι συμβατές. Η [Wu et al., 2004] προτείνει μεθόδους επίτευξης συμβατής εσωτερικής κατάστασης με την ελάχιστη δυνατή επιβάρυνση της κανονικής λειτουργίας των εξυπηρετητών.

Στην [Grefen et al., 2001] περιγράφεται φορμαλιστικά η επέκταση στο πρωτόκολλο συναλλαγών Saga [Garcia-Molina and Salem, 1987] για τη διαχείριση ροής εργασιών. Οι συναλλαγές διακρίνονται σε τοπικές, όπου ισχύουν οι περιορισμοί και οι εγγυήσεις ACID των συναλλαγών, και σε κατανεμημένες συναλλαγές που συνθέτουν τις τοπικές σε ροές εργασιών. Επεξηγείται αναλυτικά ο αλγόριθμος για τον προσδιορισμό της ορθότητας ροών εργασίας με βάση το γράφο εκτέλεσης της ροής και των επιμέρους τοπικών συναλλαγών.

Η σημαντικότερη συνεισφορά της εργασίας είναι η φορμαλιστική περιγραφή του προτεινόμενου μοντέλου συναλλαγών και η υλοποίηση και ενσωμάτωση του μοντέλου σε μία αρχιτεκτονική συστήματος.

Στην [Samaras et al., 1993] εξετάζονται το βασικό 2PC πρωτόκολλο και 10 παραλλαγές του. Υπολογίζονται ενέργειες που χαρακτηρίζονται από υψηλό κόστος όπως η επιβεβλημένη άμεση καταγραφή πληροφοριών στο ιστορικό, η ανταλλαγή μηνυμάτων και το κόστος λάθος επιλογών στα πρωτόκολλα με ευρετικές βελτιώσεις, με απώτερο στόχο την εκτίμηση της απόδοσης σε διάφορες συνθήκες λειτουργίας. Ζητήματα που αφορούν το κόστος αποκλεισμού στην περίπτωση προσωρινής απώλειας επικοινωνίας με το διαχειριστή και ο ανταγωνισμός για διαμοιρασμένους πόρους δεν εξετάζονται.

Η απόδοση των πρωτοκόλλων κατανεμημένων συναλλαγών ερευνάται επίσης στην [Gupta et al., 1996] συγκριτικά με παραλλαγές του βασικού πρωτοκόλλου ειδικά προσαρμοσμένες για περιβάλλοντα πραγματικού χρόνου με αυστηρές απαιτήσεις για το χρόνο ολοκλήρωσης της συναλλαγής. Το μοντέλο προσομοίωσης περιλαμβάνει το διαχειριστή συναλλαγών, το διαχειριστή προσπέλασης σε διαμοιρασμένους πόρους, το διαχειριστή πόρων και το διαχειριστή ανάνηψης από αποτυχίες πτώσης.

Η περίπτωση αποτυχιών πτώσης του εξυπηρετητή και η επακόλουθη επίπτωση σε μετρικές απόδοσης, ερευνάται στη [Liu et al., 1998]. Τα συμπεράσματα εξάγονται από την επεξεργασία δεδομένων προσομοίωσης για το βασικό πρωτόκολλο 2PC και τις δύο δημοφιλέστερες εκδοχές του, τα πρωτόκολλα 2PC PRA και το 2PC PRC. Ιδιαίτερο ενδιαφέρον παρουσιάζει η μελέτη των τιμών των παραμέτρων για τις οποίες η συμπεριφορά του πρωτοκόλλου αλλάζει ραγδαία (thrashing point). Εξετάζονται φόρτοι εργασίας σε συνθήκες όπου εμφανίζονται αποτυχίες πτώσης και σε ιδανικές συνθήκες χωρίς αποτυχίες, για να μελετηθεί η επίπτωση της διαδικασίας ανάνηψης στις μετρικές απόδοσης.

Σε ειδικά περιβάλλοντα εκτέλεσης, όπως οι κινητές συσκευές υπολογισμού, οι διαφορετικές προσδοκίες των χρηστών και οι περιορισμένες δυνατότητες των συσκευών σε υπολογιστική ισχύ και διαθέσιμη μνήμη επέβαλλαν τον επανασχεδιασμό του αρχικού πρωτοκόλλου 2PC ώστε αυτό να λειτουργεί αποδοτικότερα. Οι επιπτώσεις των σχεδιαστικών επιλο-

γών στην απόδοση των συστημάτων επεξεργασίας συναλλαγών κινητών συσκευών, αποτιμώνται στην [Nouali-Taboudjemat et al., 2010]. Παράγοντες, όπως το κόστος επικοινωνίας σε κυψελοειδή δίκτυα, αλλά και όσα απαντώνται σε κλασσικά πρωτόκολλα συναλλαγών, μελετώνται με προσομοίωση για να εκτιμηθεί η απόδοση που επιτυγχάνουν τα εξεταζόμενα πρωτόκολλα. Προτείνουν τη δική τους παραλλαγή του 2PC, το aTCP, που ορίζεται από διάγραμμα καταστάσεων–μεταβάσεων. Δε μελετώνται σχέσεις αντιστάθμισης μεταξύ των μετρικών απόδοσης ούτε ελέγχονται οι ιδιότητες του μοντέλου.

Σε σχέση με τις εργασίες που αναφέρθηκαν, η παρούσα διατριβή αποσκοπεί στη μελέτη των συνδυασμένων επιδράσεων των διαφόρων πρωτοκόλλων στο πλαίσιο μιας αρχιτεκτονικής επεξεργασίας συναλλαγών. Ενώ στη σχετική βιβλιογραφία εξετάζονται επιμέρους πτυχές της επεξεργασίας συναλλαγών, η κάθε μία ανεξάρτητα, στη διατριβή αυτή εξετάζουμε τον τρόπο με τον οποίο οι πτυχές αυτές αλληλεπιδρούν και τελικά καθορίζουν την διακύμανση της απόδοσης και της διαθεσιμότητας επεξεργασίας συναλλαγών. Άλλη μία διαφορά είναι ότι ενώ οι σχετικές εργασίες αναφέρονται σε συναλλαγές δεδομένων, στη δική μας περίπτωση επικεντρωνόμαστε σε προγράμματα που εκτελούνται ως κατανεμημένες συναλλαγές σε πληροφοριακό σύστημα. Κατά συνέπεια, αναφερόμαστε σε συναλλαγές που παρέχουν εγγυήσεις για διαμοιρασμένες δομές δεδομένων με μόνιμη αποθήκευση (persistent state).

## 2.3 Τυπική ανάλυση επεξεργασίας κατανεμημένων συναλλαγών

Υπάρχουν αρκετές προτάσεις στη βιβλιογραφία για τον ορισμό μοντέλων (προηγμένων) συναλλαγών και για τον έλεγχο των ιδιοτήτων τους. Ωστόσο, δε γνωρίζουμε κάποια εργασία που να συνδυάζει την επικύρωση μοντέλων με την αυτόματη σύνθεση κώδικα από το ελεγμένο μοντέλο. Οι περισσότερες προσεγγίσεις υποθέτουν την εξοικείωση του σχεδιαστή με μία τυπική γλώσσα στην οποία στηρίζεται η προσέγγιση που ακολουθούν. Επιπλέον, η επι-

κύρωση του μοντέλου πραγματοποιείται με την απόδειξη της ορθότητας των εξεταζόμενων ιδιοτήτων, αντί της περιορισμένου κόστους αλγοριθμικής προσέγγισης που υιοθετήσαμε στο Κεφάλαιο 3. Μία άλλη διαφορά είναι ότι τα μοντέλα που χρησιμοποιούμε στηρίζονται σε αυτόματα και δεν απαιτείται εξοικείωση με κάποια συγκεκριμένη τυπική γλώσσα.

Στη [Lynch and Merritt, 1986], η συμπεριφορά των οντοτήτων που συμμετέχουν στην επεξεργασία ενθυλακωμένων συναλλαγών ορίζεται από δίκτυο αυτομάτων I/O. Οι ιδιότητες ορθότητας επαληθεύονται με μαθηματικές αποδείξεις. Το άρθρο επικεντρώνεται σε ιδιότητες που αφορούν την ισοδυναμία με σειριακή εκτέλεση των συναλλαγών και τον έλεγχο ταυτοχρονισμού για την προσπέλαση σε διαμοιρασμένους πόρους. Η δική μας έρευνα επικεντρώνεται σε διαφορετικά ερωτήματα που όμως είναι συμπληρωματικά σε αυτά που αναφέρθηκαν.

Η εργασία [Johnson et al., 2007] επικεντρώνεται σε συναλλαγές υπηρεσιών ιστού. Προτείνει το πρωτόκολλο συναλλαγών WS-AT που επεκτείνει το βασικό πρωτόκολλο 2PC με επιπλέον ιδιότητες. Αξιοποιεί τη Χρονική Λογική Ενεργειών (TLA) για την μελέτη του προτεινόμενου πρωτοκόλλου. Οι ιδιότητες ορθότητας επαληθεύονται με τον ελεγκτή μοντέλων TLC και εκφράζονται ως αμετάβλητες λογικές προτάσεις που χαρακτηρίζουν τις καταστάσεις του μοντέλου. Στόχος της εργασίας είναι ο ακριβής ορισμός του μοντέλου ώστε να αποφεύγονται ασάφειες αλλά δεν αντιμετωπίζονται ζητήματα σχετικά με την ορθή υλοποίηση της επεξεργασίας συναλλαγών.

Το ACTA [Chrysanthis and Ramamritham, 1990] είναι ένα μαθηματικό πλαίσιο για τον ορισμό και τον έλεγχο των επιπτώσεων στα αντικείμενα που συμμετέχουν σε συναλλαγές και των μεταξύ τους αλληλεπιδράσεων. Οι υποθέσεις σχετικά με τα αντικείμενα και το ιστορικό των πιθανών επιδράσεων σ' αυτά, εκφράζονται με αξιώματα. Τα θεωρήματα αντιστοιχούν σε ιδιότητες του μοντέλου και εκφράζουν πάγιες σχέσεις μεταξύ των αντικειμένων που αποδεικνύονται με βάση τα διατυπωμένα αξιώματα. Καινούρια μοντέλα συναλλαγών μπορούν να οριστούν αλλάζοντας κάποιο από τα αξιώματα ή με την εισαγωγή νέων σχέσεων μεταξύ των αντικειμένων. Για τον έλεγχο των ιδιοτήτων στο καινούριο μοντέλο αρκεί να αποδειχθούν τα θεωρήματα που βασίζονται σε αξιώματα που έχουν μεταβληθεί στο νέο

μοντέλο. Όσα βασίζονται σε αξιώματα που δεν άλλαξαν, ισχύουν και στο καινούριο μοντέλο. Ενδεικτικά, στην [Chrysanthis and Ramamritham, 1994] ορίζονται και μελετώνται οι ιδιότητες των ενθυλακομένων συναλλαγών, αλλά και άλλων μοντέλων προηγμένων συναλλαγών, επεκτείνοντας τα αξιώματα του βασικού μοντέλου συναλλαγών. Το ACTA είναι εξαιρετικά αποτελεσματικό στον ορισμό και το φορμαλιστικό έλεγχο ιδιοτήτων νέων μοντέλων συναλλαγών, αλλά απαιτεί καλή γνώση της μαθηματικής λογικής και της θεωρίας συνόλων. Επιπλέον, το ACTA δεν εξετάζει ζητήματα που αφορούν την υλοποίηση του μοντέλου συναλλαγών.

Από τις παραπάνω εργασίες, η δική μας προσέγγιση μοιάζει περισσότερο με εκείνη που στηρίζεται στο TLA η οποία αντιμετωπίζει και ζητήματα σχετικά με τον έλεγχο ιδιοτήτων ασφάλειας (safety). Ωστόσο, με τη δική μας προσέγγιση έχουμε επιπλέον τη δυνατότητα ελέγχου ιδιοτήτων βιωσιμότητας (liveness) και προσεγγισιμότητας (reachability). Αν υπάρχει η δυνατότητα χρήσης φορμαλιστικών μεθόδων μέσω αποδεικτικής προσέγγισης και δε μας απασχολούν ζητήματα υλοποίησης του μοντέλου, το ACTA προσφέρει τις περισσότερες δυνατότητες στη διατύπωση και τον έλεγχο ιδιοτήτων των συναλλαγών.

## **Κεφάλαιο 3**

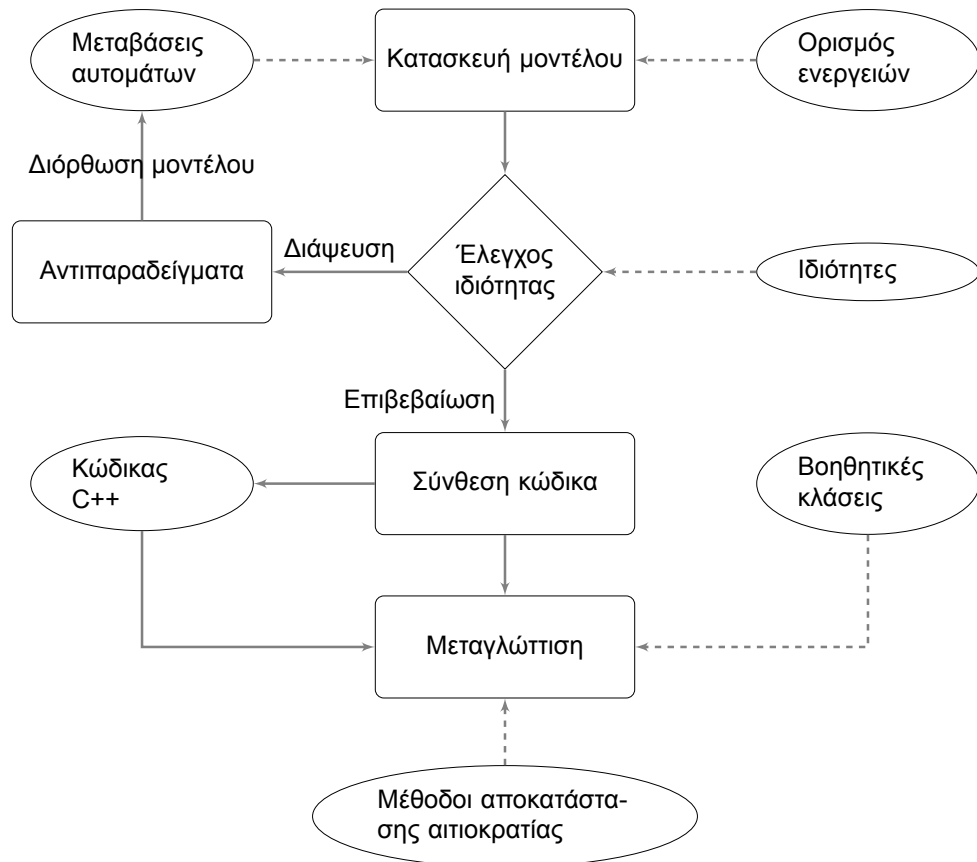
# **Έλεγχος μοντέλων και σύνθεση κώδικα επεξεργασίας κατανεμημένων συναλλαγών**

Τα μοντέλα συναλλαγών περιγράφουν εργασίες που πρόκειται να εκτελεστούν από συναλλασσόμενες οντότητες με σαφώς καθορισμένους ρόλους. Οι οντότητες επικοινωνούν με ασύγχρονα μηνύματα. Με τη λήψη ενός μηνύματος, κάθε ρόλος διεκπεραιώνει υπολογισμούς που ορίζονται από το πρωτόκολλο συναλλαγών. Η υλοποίηση συστημάτων επεξεργασίας συναλλαγών είναι δύσκολη και επίπονη διαδικασία λόγω των πληροφοριών που πρέπει να χειριστεί το σύστημα για τη σωστή δέσμευση και απελευθέρωση πόρων, ώστε να εξυπηρετούνται οι ανάγκες των ενεργών συναλλαγών. Σε αυτό το κεφάλαιο παρουσιάζουμε μία προσέγγιση ανάπτυξης συστημάτων επεξεργασίας συναλλαγών καθοδηγούμενη από μοντέλα πεπερασμένων καταστάσεων [Stahl et al., 2006]. Υποστηρίζει έλεγχο των επιδιωκόμενων ιδιοτήτων [Clarke et al., 1999] των μοντέλων συναλλαγών και αυτόματη σύνθεση μέρους της υλοποίησης από το εξακριβωμένα ορθό μοντέλο. Σχετικές εργασίες [Chrysanthis and Ramamritham, 1994, Johnson et al., 2007, Lynch and Merritt, 1986]

προϋποθέτουν την ύπαρξη ειδικού φορμαλισμού για τον καθορισμό και τη μετέπειτα ανάλυση μοντέλων. Στη δική μας προσέγγιση, οι οντότητες του μοντέλου συναλλαγών ορίζονται με εύληπτο τρόπο από μη αιτιοκρατικά αυτόματα. Τα μηνύματα που ανταλλάσσονται μεταξύ των οντοτήτων του μοντέλου στέλνονται από τις ενέργειες που εκτελούνται κατά τις μεταβάσεις των αυτομάτων. Οι ιδιότητες ορθότητας ελέγχονται σε όλα τα πιθανά μονοπάτια εκτέλεσης του σύνθετου αυτόματου που προκύπτει από το συγχρονισμένο γινόμενο των επιμέρους αυτομάτων. Ο έλεγχος του μοντέλου δεν απαιτεί γνώσεις τυπικών μεθόδων. Το επικυρωμένο μοντέλο χρησιμεύει για τη σύνθεση κώδικα που χειρίζεται τα μηνύματα των μοντέλων συναλλαγών. Τα παραπάνω χαρακτηριστικά διακρίνουν τη δική μας μέθοδο από άλλες προσεγγίσεις που βασίζονται σε ελεγκτές μοντέλων γενικής χρήσης όπως το SPIN [Holzmann, 2003]. Η δική μας προσέγγιση δανείζεται ιδέες από τους παραγωγούς μεταγλωττιστών όπου ένα εργαλείο ελέγχει τη συντακτική προδιαγραφή της γλώσσας και παράγει την υλοποίηση του συντακτικού αναλυτή που ενσωματώνει και κώδικα που παρέχει ο χρήστης.

Ο έλεγχος ιδιοτήτων και η σύνθεση κώδικα έχουν υλοποιηθεί στο εργαλείο ACID Model Checker & Code Generator [Mentis and Katsaros, 2009]. Το πρόγραμμα παράγει κώδικα για το ACID Sim Tools [Mentis et al., 2008], έναν προσομοιωτή για τη μελέτη σημείων αντιστάθμισης της απόδοσης και του χρόνου ανάνηψης από αποτυχία πτώσης σε συστήματα επεξεργασίας συναλλαγών [Katsaros et al., 2007]. Ωστόσο, η συνεισφορά της παρούσας εργασίας μπορεί να αξιοποιηθεί και στην ανάπτυξη συστημάτων επεξεργασίας συναλλαγών. Θεωρούμε ότι η προτεινόμενη μέθοδος έχει ευρύτερες εφαρμογές από το ACID Sim Tools.

Σε αυτό το κεφάλαιο παρουσιάζουμε τα αποτελέσματα που προέκυψαν από τις εργασίες [Mentis and Katsaros, 2009] και [Mentis and Katsaros, 2012a]. Στην πρώτη εργασία παρουσιάσαμε αποτελέσματα για το βασικό μοντέλο συναλλαγών 2PC ενώ στη δεύτερη εξετάσαμε το μοντέλο ενθυλακωμένων συναλλαγών [Moss, 1982], ένα παράδειγμα προηγμένου μοντέλου κατανεμημένων συναλλαγών. Στις ενθυλακωμένες συναλλαγές υπάρχουν επιπλέον οντότητες με αποτέλεσμα τη σημαντική αύξηση του χώρου καταστάσεων. Για την



Σχήμα 3.1: Επιβεβαίωση και υλοποίηση μοντέλων συναλλαγών

αντιμετώπιση του μεγαλύτερου χώρου καταστάσεων προσθέσαμε ένα στάδιο προεργασίας στον αρχικό αλγόριθμο που αφαιρεί τις μεταβάσεις που δεν επηρεάζουν τις εξεταζόμενες ιδιότητες. Επιπλέον, υπάρχει η δυνατότητα καθορισμού συνθηκών απενεργοποίησης των μεταβάσεων που εισάγουν μη αιτιοκρατική συμπεριφορά στο μοντέλο.



### 3.1 Ανάπτυξη συστημάτων επεξεργασίας συναλλαγών καθοδηγούμενη από μοντέλα πεπερασμένων καταστάσεων

Η διαδικασία για την εξακρίβωση ιδιοτήτων μοντέλου συναλλαγών και την επακόλουθη αυτόματη σύνθεση κώδικα για το χειρισμό μηνυμάτων (Σχήμα 3.1) συμπεριλαμβάνει τον ορισμό του μοντέλου, τις αναμενόμενες ιδιότητές του και την εξερεύνηση των εφικτών σεναρίων εκτέλεσης ώστε να εξακριβωθεί η εγκυρότητα των ιδιοτήτων.

Τα μοντέλα ορίζονται από τα αυτόματα των αλληλεπιδρώντων οντοτήτων και από τα μηνύματα που αποστέλλονται, προγραμματίζονται για αποστολή ή ακυρώνονται από ενέργειες που εκτελούνται καθώς τα αυτόματα μεταβαίνουν σε νέες καταστάσεις. Κατά την κατασκευή του μοντέλου υπολογίζεται το γράφημα προσβασιμότητας του συγχρονισμένου γινόμενου [Bérard et al., 2001] των αυτομάτων των οντοτήτων. Το συγχρονισμένο γινόμενο προκύπτει από το καρτεσιανό γινόμενο των αυτομάτων περιορισμένο από τα μηνύματα συγχρονισμού, εκείνα δηλαδή τα μηνύματα όπου ο παραλήπτης είναι διαφορετική οντότητα από τον αποστολέα. Τα μονοπάτια στο γράφημα προσβασιμότητας αντιστοιχούν σε διαφορετικούς συνδυασμούς σειράς παραλαβής μηνυμάτων και χρησιμοποιούνται για τη διαπίστωση λαθών στον ορισμό του μοντέλου (όπως απροσπέλαστες καταστάσεις και μερικώς ορισμένη συνάρτηση μετάβασης) και την εύρεση εκτελέσεων όπου παραβιάζουν τις εξεταζόμενες ιδιότητες. Τα μονοπάτια όπου διαπιστώνεται παραβίαση κάποιας ιδιότητας, αναφέρονται στο χρήστη ως αντιπαραδείγματα της εγκυρότητας της ιδιότητας. Τα αντιπαραδείγματα βοηθούν στη διόρθωση του μοντέλου.

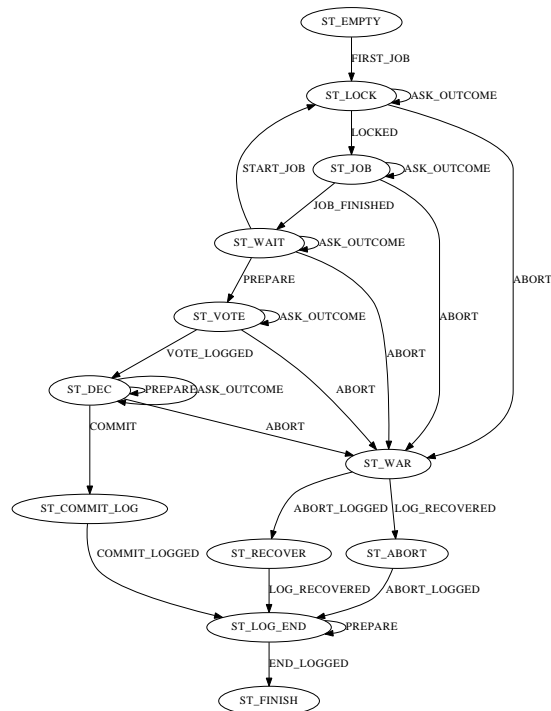
Ο κώδικας που παράγεται από το μοντέλο αφορά τη μέθοδο που χειρίζεται τις μεταβάσεις των καθορισμένων ρόλων. Η παραγόμενη μέθοδος προσδιορίζει την επόμενη κατάσταση του ρόλου – παραλήπτη του μηνύματος και καλεί τις ενέργειες που σχετίζονται με την ενεργοποιημένη μετάβαση. Η επιλογή της μετάβασης γίνεται με βάση το τύπο του μηνύματος και την τρέχουσα κατάσταση του παραλήπτη. Όταν προκύπτουν παραπάνω από

μία επιλογή, ο χειρισμός του μηνύματος δεν είναι αιτιοκρατικός. Η μη αιτιοκρατική συμπεριφορά απορρέει από την αβεβαιότητα και την ανεπαρκή γνώση σχετικά με πληροφορίες που είναι διαθέσιμες μόνο κατά την εκτέλεση της συναλλαγής. Στη κατασκευή του μοντέλου, η έλλειψη αιτιοκρατίας αντιμετωπίζεται με τη διερεύνηση όλων των πιθανών σεναρίων εκτέλεσης. Η παραγόμενη μέθοδος ενσωματώνει κώδικα που παρέχει ο χρήστης για την αποκατάσταση της αιτιοκρατίας κατά της εκτέλεση, όταν θα είναι πλέον διαθέσιμες όλες οι απαραίτητες πληροφορίες για να αποφασιστεί η επόμενη κατάσταση. Ο κώδικας προσδιορίζει την επόμενη κατάσταση από τις διαθέσιμες επιλογές. Αν από λάθος στην υλοποίηση η επιλεγμένη κατάσταση δεν ανήκει στο σύνολο των δυνατών επιλογών, η εκτέλεση σταματά αφού πρώτα αναφερθεί το σφάλμα.

## 3.2 Το μοντέλο συναλλαγών δύο φάσεων (2PC)

Ένα κοινό γνώρισμα των κατανεμημένων μοντέλων συναλλαγών είναι ότι κάποια διεργασία (ο συντονιστής) και ομάδα από άλλες διεργασίες (οι μέτοχοι) συνεργάζονται για τη διεκπεραίωση των ενεργειών της συναλλαγής. Με βάση αυτό το μοντέλο, έχουν προταθεί αρκετά πρωτόκολλα συναλλαγών που έχουν αφετηρία το βασικό πρωτόκολλο δύο φάσεων (2PC) [Lampson and Sturgis, 1979]. Σε αυτή την ενότητα περιγράφουμε συνοπτικά το βασικό πρωτόκολλο 2PC και δύο δημοφιλείς παραλλαγές του. Στη [Mohan et al., 1986] μπορεί ο ενδιαφερόμενος αναγνώστης να βρει εκτενέστερη περιγραφή των παραπάνω πρωτοκόλλων.

Οι επιμέρους εργασίες κατανέμονται στους μετόχους της συναλλαγής. Στην πρώτη φάση του πρωτοκόλλου ο συντονιστής στέλνει στους μετόχους αίτημα διεκπεραίωσης των ενεργειών. Υποθέτουμε ότι υπάρχει μηχανισμός που παρέχει τη δυνατότητα επαναφοράς των δεδομένων στην αρχική τους κατάσταση σε περίπτωση ακύρωσης της συναλλαγής. Συνήθως, οι ενέργειες που πρόκειται να εκτελεστούν και τα δεδομένα που πρόκειται να μεταβληθούν από τις ενέργειες καταχωρούνται σε ιστορικό που διατηρείται σε μέσο μόνιμης



Σχήμα 3.2: Απλοποιημένα διαγράμματα του αυτόματου ρόλου μετόχου στο βασικό πρωτόκολλο συναλλαγών δύο φάσεων

αποθήκευσης. Αν ακυρωθεί η συναλλαγή, τα αρχικά δεδομένα αποκαθιστώνται με βάση το ιστορικό. Μία άλλη προϋπόθεση είναι ότι υπάρχει η δυνατότητα επιβεβλημένης καταγραφής δεδομένων σε μέσο μόνιμης αποθήκευσης. Τα δεδομένα για μόνιμη αποθήκευση μπορεί να διατηρούνται προσωρινά στην κεντρική μνήμη από το λειτουργικό σύστημα για λόγους απόδοσης. Σε περίπτωση όμως αποτυχίας πτώσης συστήματος, τα δεδομένα χάνονται καθιστώντας αδύνατη την επαναφορά του συστήματος. Με τη λήψη του αιτήματος προετοιμασίας για οριστικοποίηση (PREPARE), ο μέτοχος εκτελεί επιβεβλημένη εγγραφή στο ιστορικό ώστε να γνωρίζει την ύπαρξη της συναλλαγής σε περίπτωση ανάνηψης από αποτυχία του εξυπηρετητή.

Οι μεταβάσεις του ρόλου του μετόχου συναλλαγών απεικονίζονται στο Σχήμα 3.2. Οι ακμές αντιστοιχούν σε μεταβάσεις του αυτομάτου ενώ ο τύπος του μηνύματος που την προκαλεί αναγράφεται πάνω στο τόξο. Ορισμένες μεταβάσεις που αφορούν ετεροχρονισμένη

λήψη μηνυμάτων αγνοούνται από το ρόλο και δεν εμφανίζονται.

Ο μέτοχος αποκρίνεται θετικά στο αίτημα του συντονιστή αν εκτελέσει με επιτυχία όλες τις ενέργειες της συναλλαγής που του αντιστοιχούν. Σε περίπτωση σφάλματος κατά την εκτέλεση κάποιας ενέργειας ή αν δεν είναι διαθέσιμοι οι πόροι για τη διεκπεραίωση των ενεργειών (οι πόροι έχουν δεσμευτεί από άλλες συναλλαγές, λόγω χάρη), ο μέτοχος αποφασίζει να ακυρώσει τη συναλλαγή και εκτελεί επιβεβλημένη καταχώριση της πρόθεσης του στο ιστορικό απαντά αρνητικά. Οι μέτοχοι της συναλλαγής έχουν δικαίωμα βέτο στην οριστικοποίηση της συναλλαγής αλλά δεσμεύονται αν απαντήσουν θετικά. Τα αποτελέσματα της συναλλαγής σε κάποιο μέτοχο που την υπερψήφισε μπορούν πλέον να ακυρωθούν μόνο με απόφαση του συντονιστή. Πόροι που έχουν δεσμευτεί για τη συναλλαγή δε μπορούν να αποδεσμευθούν μέχρι να ληφθεί η απόφαση από το συντονιστή. Αν χαθεί η μεταξύ τους επικοινωνία, οι άλλες συναλλαγές αποκλείονται από τη πρόσβαση στους δεσμευμένους πόρους μέχρι την αποκατάσταση της επικοινωνίας. Γι' αυτό, το βασικό πρωτόκολλο 2PC και οι παραλλαγές του αποκαλούνται πρωτόκολλα αποκλεισμού.

Ο ρόλος του συντονιστή στην πρώτη φάση επεξεργασίας της συναλλαγής είναι ο συντονισμός της διαδικασίας ψήφου. Αφού σταλεί στους μετόχους το αίτημα προετοιμασίας της συναλλαγής, ο συντονιστής συλλέγει τις απαντήσεις τους και ενημερώνει τους μετόχους για τη τελική έκβαση. Αν δεν ασκηθεί βέτο, ο συντονιστής καταχωρεί στο ιστορικό επιβεβλημένη εγγραφή που σηματοδοτεί την πρόθεση για οριστικοποίηση της συναλλαγής και ενημερώνει τους μετόχους για την απόφαση. Αντίστοιχα πράττει και σε περίπτωση άρνησης κάποιου μετόχου, με τη διαφορά ότι η καταχώριση στο ιστορικό σηματοδοτεί την ακύρωση της συναλλαγής. Οι μέτοχοι με τη σειρά τους, καταχωρούν στο ιστορικό την απόφαση του συντονιστή και επιβεβαιώνουν την εφαρμογή της. Ο συντονιστής καταχωρεί την τελική, μη επιβεβλημένη, εγγραφή στο ιστορικό που σηματοδοτεί την ολοκλήρωση της συναλλαγής όταν ληφθούν οι επιβεβαιώσεις όλων των μετόχων.

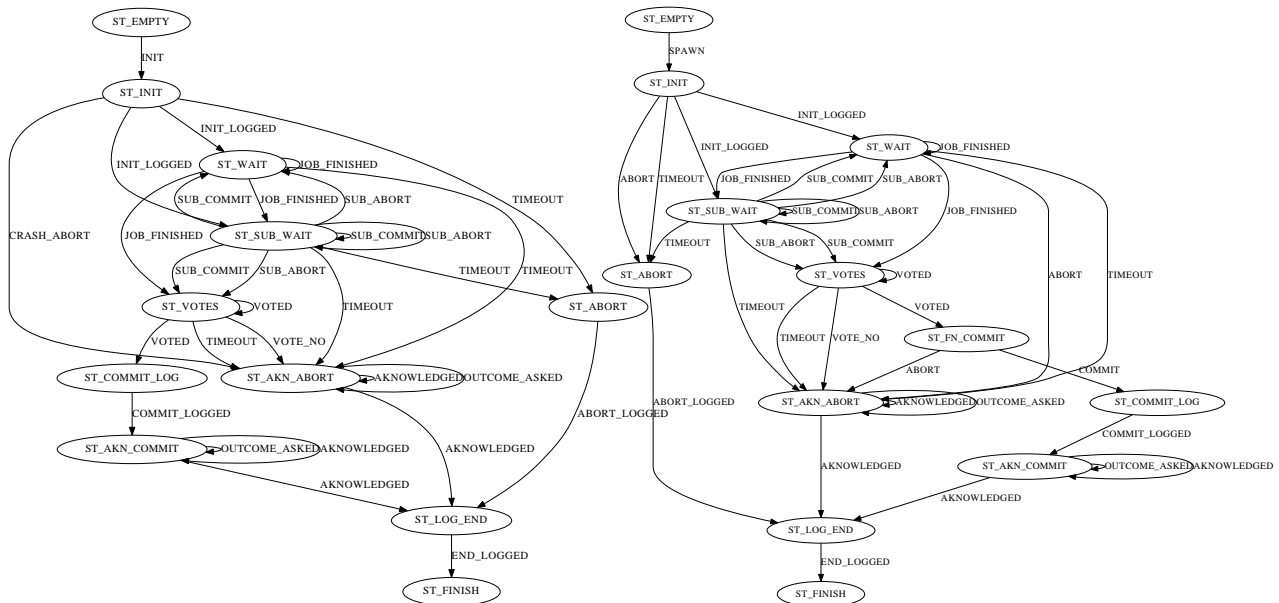
Κατά την αποκατάσταση της ομαλής λειτουργίας του εξυπηρετητή μετά την εκδήλωση αποτυχίας πτώσης, ανακτάται το ιστορικό από το μόνιμο μέσο αποθήκευσης και εξετάζονται οι καταχωρίσεις για κάθε συναλλαγή. Οι συναλλαγές για τις οποίες δεν υπάρχουν

εγγραφές στο ιστορικό και όσες έχουν καταγράψει την ολοκλήρωσή τους δε χρειάζονται αποκατάσταση και αγνοούνται. Αντίθετα, για τις συναλλαγές που δεν έχει καταγραφεί η έκβασή τους, ο μέτοχος ρωτά το συντονιστή γι' αυτές και εφαρμόζει την απόφαση που θα λάβει. Δύο παραλλαγές του βασικού 2PC υιοθετούν διαφορετική προσέγγιση σε αυτό το σημείο, με στόχο τη μείωση του αριθμού των επιβεβλημένων εγγραφών στο ιστορικό και τον περιορισμό των μηνυμάτων που πρέπει να ανταλλάγουν.

Στο 2PC PRA [Mohan et al., 1986], οι συναλλαγές των οποίων η έκβαση δεν έχει καταγραφεί στο ιστορικό θεωρούνται ακυρωμένες. Αν η υπόθεση είναι λάθος και ζητηθεί η ψήφος του μετόχου μετά την ολοκλήρωση της επαναφοράς του εξυπηρετητή, ο μέτοχος απαντά με αρνητική ψήφο. Επιπλέον, η καταγραφή της ολοκλήρωσης των ακυρωμένων συναλλαγών καθώς και η επιβεβαίωση προς το συντονιστή την απόφασης ακύρωσης είναι πλέον περιττές. Για τις οριστικοποιημένες συναλλαγές το πρωτόκολλο έχει την ίδια συμπεριφορά με το βασικό 2PC αλλά εξοικονομεί μία επιβεβλημένη καταγραφή και ένα μήνυμα (την επιβεβαίωση της απόφασης ακύρωσης) σε κάθε μέτοχο και δύο επιβεβλημένες εγγραφές στον συντονιστή. Αντίστοιχη εξοικονόμηση επιτυγχάνεται στο πρωτόκολλο 2PC PRC [Mohan et al., 1986] για τις επιτυχημένες συναλλαγές με τη διαφορά ότι ο συντονιστής πρέπει να καταγράψει τα στοιχεία των μετόχων πριν ξεκινήσει η επεξεργασία της συναλλαγής. Αν εκδηλωθεί αποτυχία πτώσης στο συντονιστή πριν την ολοκλήρωση της συναλλαγής, τα στοιχεία των μετόχων ανακτώνται από το ιστορικό για να σταλεί η απόφαση ακύρωσης.

### 3.3 Μοντέλο ενθυλακομένων συναλλαγών

Οι ενθυλακομένες συναλλαγές [Moss, 1982] είναι το πρώτο προηγμένο μοντέλο συναλλαγών ειδικά προσαρμοσμένο στις απαιτήσεις των κατανεμημένων συστημάτων. Προηγμένα μοντέλα συναλλαγών όπως οι συναλλαγές όπου μετέχουν φορητές συσκευές υπολογισμού [Serrano-Alvarado et al., 2004] και οι συναλλαγές ιστού [Wang et al., 2008], προσαρμόζουν το μοντέλο ενθυλακομένων συναλλαγών στις δικές τους ιδιαίτερες απαιτήσεις.



(I) Αυτόματο ρόλου συντονιστή γονικής συναλλαγής (II) Αυτόματο ρόλου συντονιστή υποσυναλλαγής

Σχήμα 3.3: Απλοποιημένα διαγράμματα αυτόματων των συντονιστών ενθυλακωμένων συναλλαγών

Οι συναλλαγές αποτελούνται από υποσυναλλαγές που μπορούν με τη σειρά τους να ξεκινήσουν άλλες υποσυναλλαγές σχηματίζοντας ιεραρχική δομή απεριόριστου βάθους. Η οντότητα που ξεκινά την υποσυναλλαγή είναι υπεύθυνη για το συντονισμό της διαδικασίας λήψης συνολικής απόφασης. Αναλαμβάνει τη συλλογή των αποφάσεων των άλλων οντοτήτων που συμμετέχουν στην εκτέλεση της υποσυναλλαγής και κοινοποιεί το αποτέλεσμα της ψηφοφορίας σε όλα τα εμπλεκόμενα μέλη. Η συναλλαγή που ξεκινά την υποσυναλλαγή ονομάζεται γονική ενώ η υποσυναλλαγή αναφέρεται ως παιδί της. Η γονική συναλλαγή δε μπορεί να οριστικοποιήσει τα αποτελέσματά της πριν ολοκληρωθεί η εκτέλεση όλων των παιδιών της. Αν ακυρωθεί η γονική συναλλαγή ακυρώνονται και τα παιδιά της. Ωστόσο, οι υποσυναλλαγές μπορούν να ακυρωθούν ανεξάρτητα από την έκβαση της γονικής συναλλαγής. Οι μεταβάσεις των δύο ρόλων συντονιστών απεικονίζονται στο Σχήμα 3.3. Όπως φαίνεται στο Σχήμα 3.3, ο συντονιστής υποσυναλλαγών δεν οριστικοποιεί τα αποτελέσματα της συναλλαγής μετά την ολοκλήρωση της ψηφοφορίας αν δε λάβει την απόφαση (COM-

MIT) του συντονιστή της γονικής συναλλαγής.

### 3.4 Ορισμός ρόλων

Τα μοντέλα συναλλαγών ορίζονται με βάση τους ρόλους που επιτελούν οι εμπλεκόμενες οντότητες όπου κάθε ρόλος περιγράφεται από ένα μη αιτιοκρατικό αυτόματο. Στο πρωτόκολλο 2PC για παράδειγμα, υπάρχουν δύο ρόλοι, ο συντονιστής της συναλλαγής και οι απλοί μέτοχοι σε αυτή. Το ενθυλακωμένο μοντέλο έχει τέσσερις ρόλους, ο συντονιστής και ο μέτοχος σε συναλλαγές καθώς και οι αντίστοιχοι ρόλοι υποσυναλλαγών. Τα προηγμένα μοντέλα συναλλαγών έχουν γενικά περισσότερους από δύο ρόλους.

Η έλλειψη αιτιοκρατίας διευκολύνει στον ορισμό των μοντέλων και γι' αυτό είναι αποδεκτή και από τις γλώσσες άλλων δημοφιλών ελεγκτών μοντέλων [Holzmann, 2003]. Πέρα από τη διευκόλυνση στην έκφραση των μοντέλων, η μη αιτιοκρατία είναι απαραίτητη όταν η συμπεριφορά του ρόλου εξαρτάται από πληροφορίες που είναι διαθέσιμες μόνο κατά την εκτέλεση. Για παράδειγμα, όταν ο συντονιστής λάβει στο στάδιο της ψηφοφορίας την απόφαση ενός μετόχου, είτε κοινοποιεί την απόφαση σε όλους τους μετόχους αν έλαβε την τελευταία θετική ψήφο ή αίτημα ακύρωσης της συναλλαγής, είτε περιμένει τους υπόλοιπους μετόχους να ψηφίσουν. Η έλλειψη αιτιοκρατίας είναι αναπόφευκτη αφού η συμπεριφορά του συντονιστή για την ίδια κατάσταση (αναμονή ψήφου) και το ίδιο μήνυμα (θετική ψήφος) διαφοροποιείται ανάλογα με το αν πρόκειται για την τελευταία αναμενόμενη ψήφο. Η μέθοδος αποκατάστασης της αιτιοκρατίας στον έλεγχο μοντέλων παρουσιάζεται στην Ενότητα 3.6.

Το αλφάβητο των αυτομάτων των ρόλων περιλαμβάνει όλα τα μηνύματα που μπορεί να ληφθούν κατά την επεξεργασία συναλλαγών. Τα μηνύματα προκαλούν τη μετάβαση σε νέα κατάσταση του αυτομάτου του ρόλου – παραλήπτη. Επιπλέον, καλούνται οι ενέργειες που σχετίζονται με την εκτελούμενη μετάβαση όπως ορίζει η περιγραφή του μοντέλου. Οι καλούμενες ενέργειες μπορεί: (i) να στείλουν μηνύματα (ii) να προγραμματίσουν την αποστολή μηνυμάτων (iii) να ακυρώσουν κάποιες από τις προγραμματισμένες αποστολές μη-

νυμάτων (iv) ή απλά να εκτελέσουν υπολογισμούς. Η σχέση μετάβασης των ρόλων του μοντέλου ορίζεται από αρχείο κειμένου με πέντε (5) στήλες διαχωρισμένες με κόμμα.

**Ρόλος** Η πρώτη στήλη ορίζει το όνομα του ρόλου στον οποίο αντιστοιχεί η μετάβαση του αυτομάτου. Στο βασικό μοντέλο συναλλαγών έχουμε τα ονόματα “c” και “w” για το συντονιστή (coordinator) και μέτοχο (worker) αντίστοιχα. Στο ενθυλακωμένο μοντέλο συναντάμε άλλους δύο ρόλους (“cn” και “wn”) για το συντονιστή και το μέτοχο υπο-συναλλαγών.

**Αρχική κατάσταση** Η κατάσταση στην οποία επιτρέπεται η μετάβαση.

**Μήνυμα** Το μήνυμα που ενεργοποιεί τη μετάβαση. Το μήνυμα προέρχεται από το σύνολο μηνυμάτων (αλφάβητο) που αναγνωρίζει ο παραλήπτης.

**Επόμενη κατάσταση** Η κατάσταση στην οποία μεταβαίνει το αυτόματο με τη λήψη του μηνύματος.

**Ενέργειες** Ονόματα μεθόδων που καλούνται κατά τη μετάβαση. Τα ονόματα διαχωρίζονται από το σύμβολο “:”. Οι μεταβάσεις που δεν εκτελούν ενέργειας, χρησιμοποιούν το σύμβολο “-” για να σηματοδοτήσουν την απουσία ενεργειών. Κάθε ενέργεια έχει ως πρώτη παράμετρο το προσδιοριστικό της συναλλαγής. Το όνομα των ενεργειών που δέχονται επιπλέον παραμέτρους ακολουθείται από σειρά προσδιοριστικών που περικλείονται σε άγκιστρα. Τα προσδιοριστικά αντιστοιχούν σε παραμέτρους της μεθόδου όπως ο παραλήπτης του μηνύματος και η τρέχουσα εργασία της συναλλαγής.

Η πρώτη γραμμή στο Σχήμα 3.4 ορίζει τη μετάβαση από την κατάσταση ST\_EMPTY που είναι η αρχική κατάσταση για όλους τους ρόλους. Η μετάβαση ενεργοποιείται από το μήνυμα INIT που είναι το μόνο μήνυμα που δε προκαλείται από την εκτέλεση κάποιας ενέργειας αλλά σηματοδοτεί την άφιξη της συναλλαγής στο σύστημα για επεξεργασία. Ο συντονιστής μεταβαίνει στην κατάσταση ST\_INIT και εκτελεί την ενέργεια sendInitLog η οποία παραθέτει μία εγγραφή στο ιστορικό που χρησιμεύει στη αποκατάσταση της εκτέλεσης της



	Ρόλος	Αρχική κατάσταση	Μήνυμα	Επόμενη κατάσταση	Ενέργειες
1	c,	ST_EMPTY,	INIT,	ST_INIT,	sendInitLog:scheduleTimeout
2	cn,	ST_EMPTY,	SPAWN,	ST_INIT,	sendInitLog:scheduleTimeout
3	c,	ST_VOTES,	OUTCOME_ASKED,	ST_VOTES,	-
4	c,	ST_VOTES,	VOTED,	ST_VOTES,	collectVote{sender}
5	w,	ST_WAIT,	START_JOB,	ST_LOCK,	workerLock{job}
6	w,	ST_COMMIT_LOG,	COMMIT_LOGGED,	ST_LOG_END,	clearLocks:acknowledgeCommit
7	c,	ST_WAIT,	JOB_FINISHED,	ST_WAIT,	nextJob:moveNext
8	c,	ST_WAIT,	JOB_FINISHED,	ST_VOTES,	sendPrepare

Σχήμα 3.4: Τμήμα προδιαγραφής ρόλων του μοντέλου ενθυλακομένων συναλλαγών

συναλλαγής σε περίπτωση αποτυχίας του εξυπηρετητή ή αν αποφασιστεί η ακύρωσή της. Όταν ολοκληρωθεί η καταγραφή, ο συντονιστής λαμβάνει το μήνυμα INIT\_LOGGED που δε συμπεριλαμβάνεται στο απόκομμα. Η άλλη ενέργεια που εκτελείται κατά τη μετάβαση, η `scheduleTimeout`, προγραμματίζει την λήψη του μηνύματος TIMEOUT μετά την πάροδο κάποιου χρόνου. Το μήνυμα ακυρώνεται με την ολοκλήρωση της συναλλαγής με αποτέλεσμα να φτάνει στο συντονιστή μόνο εάν η συναλλαγή δεν ολοκληρωθεί έγκαιρα. Η δεύτερη γραμμή ορίζει μετάβαση του συντονιστή υποσυναλλαγών. Αντίθετα από τις συναλλαγές, οι υποσυναλλαγές ξεκινούν με το μήνυμα SPAWN αντί του INIT. Η εκκίνηση υποσυναλλαγών είναι απόφαση του μοντέλου ενώ η άφιξη συναλλαγής στο σύστημα είναι απόφαση του περιβάλλοντος. Η έναρξη των δύο ειδών συναλλαγών διακρίνεται και στο ιστορικό από εγγραφές διαφορετικού τύπου.

Η τρίτη γραμμή είναι παράδειγμα μηνύματος που δεν επηρεάζει την έκβαση της συναλλαγής οπότε και αγνοείται. Λόγω της ασύγχρονης επικοινωνίας, τα μηνύματα μπορεί να ληφθούν ετεροχρονισμένα, όταν πλέον δεν έχουν σημασία και μπορούν να αγνοηθούν. Τέτοιες μεταβάσεις, παρότι δεν έχουν καμία επίπτωση στην εγκυρότητα των ιδιοτήτων του μοντέλου, πρέπει να δηλωθούν για να αποφευχθεί η παράληψη από αμέλειά άλλων σημαντικών μεταβάσεων. Μεταβάσεις που δεν έχουν δηλωθεί θεωρούνται παραλήψεις και αναφέρονται ως πιθανά σχεδιαστικά λάθη κατά των έλεγχου του μοντέλου.

Οι γραμμές 4 και 5 παρουσιάζουν ενέργειες που δέχονται επιπλέον παραμέτρους, τον αποστολέα του μηνύματος και την τρέχουσα εργασία της συναλλαγής αντίστοιχα.

Στη γραμμή 6, ο μέτοχος της συναλλαγής, αφού έχει καταγράψει την επιτυχή έκβαση

της συναλλαγής, απελευθερώνει τους πόρους που είχαν δεσμευτεί για την ολοκληρωμένη πλέον συναλλαγή με την επίκληση της ενέργειας `clearLocks`. Ο παραλήπτης του μηνύματος που παράγει η μέθοδος είναι η οντότητα που είναι υπεύθυνη για την ταυτοχρονισμένη πρόσβαση σε πόρους διαθέσιμους σε όλες τις ενεργές συναλλαγές (Lock Manager).

Οι γραμμές 7 και 8 δείχνουν τη συμπεριφορά του συντονιστή στην ολοκλήρωση εργασίας. Αν η συναλλαγή έχει επιπλέον εργασίες, ο συντονιστής τη στέλνει στον αντίστοιχο μέτοχο. Αντίθετα, αν όλες οι εργασίες έχουν ολοκληρωθεί, αρχίζει η διαδικασία ψηφοφορίας για να αποφασιστεί η έκβαση της συναλλαγής. Η απόφαση δεν είναι αιτιοκρατική αφού εξαρτάται από πληροφορίες (πλήθος εργασιών) διαθέσιμες μόνο κατά την εκτέλεση.

Ο ορισμός των ρόλων παρέχει όλες τις αναγκαίες πληροφορίες για τη σύνθεση του κώδικα που χειρίζεται τα εισερχόμενα μηνύματα. Ο παραγόμενος κώδικας και η βιβλιοθήκη με την υλοποίηση των ενεργειών που καλούνται στις μεταβάσεις, αρκούν για την παραγωγή της εκτελέσιμης αναπαράστασης του μοντέλου. Ωστόσο, οι παρεχόμενες πληροφορίες επιτρέπουν μόνο απλούς ελέγχους για την εγκυρότητα του μοντέλου όπως την ύπαρξη απροσπέλαστων καταστάσεων. Για την πλήρη έλεγχο του μοντέλου απαιτούνται επιπλέον πληροφορίες που αφορούν τη δημιουργία και την ακύρωση μηνυμάτων ως αποτέλεσμα των μεταβάσεων των αυτομάτων του μοντέλου.

### 3.5 Ορισμός μηνυμάτων συναλλαγών

Εκτός από το μήνυμα `INIT` που το στέλνει το σύστημα επεξεργασίας συναλλαγών, κάθε άλλο μήνυμα δημιουργείται από ενέργειες που εκτελούνται σε μεταβάσεις των αυτομάτων. Τα μηνύματα λοιπόν είναι απόρροια προγενέστερων μεταβάσεων. Για τον έλεγχο του μοντέλου χρειαζόμαστε τον καθορισμό των μηνυμάτων που παράγονται ή ακυρώνονται από ενέργειες των μεταβάσεων. Οι ενέργειες και τα μηνύματα που τα αφορούν ορίζονται σε αρχείο κειμένου με τέσσερις στήλες διαχωρισμένες με κόμμα:

**Ενέργεια** Το όνομα της μεθόδου χωρίς τις παραμέτρους της. Ενέργειες που καλούνται σε

μεταβάσεις πολλών ρόλων ορίζονται πολλαπλά, μία φορά για κάθε ρόλο.

**Μήνυμα** Το όνομα του μηνύματος που παράγεται ή ακυρώνεται. Όταν δεν είναι γνωστό το μήνυμα που παράγει η ενέργεια, αναγράφουμε όλα τα μηνύματα που μπορούν να παραχθούν διαχωρίζοντας τα ονόματά τους με το σύμβολο “|”. Τα μηνύματα που ακυρώνονται από την επικαλούμενη ενέργεια έχουν το πρόθεμα “-” στο ονόματά τους (-TIMEOUT).

**Παραλήπτης** Ο ρόλος του παραλήπτη του παραγόμενου μηνύματος. Το πεδίο αγνοείται σε ενέργειες που ακυρώνουν το μήνυμα.

**Αποστολέας** Το όνομα της οντότητας που παράγει ή ακυρώνει το μήνυμα. Η οντότητα είναι είτε ρόλος του μοντέλου ή κάποιο άλλο συστατικό του συστήματος. Στο βασικό μοντέλο συναλλαγών οι δύο επιπλέον οντότητες που δε περιγράφονται από ρόλους είναι η μονάδα μόνιμης αποθήκευσης πληροφοριών και η μονάδα συντονισμού ταυτόχρονης πρόσβασης σε διαμοιρασμένου πόρους.

	Ενέργεια	Μήνυμα	Παραλήπτης	Αποστολέας
1	startSubTransaction,	SPAWN,	cn,	c
2	startSubTransaction,	SPAWN,	cn,	cn
3	sendInitLog,	INIT_LOGGED,	c,	lgc
4	workerLock,	LOCKED,	w,	lcw
5	cancelTimeout,	-TIMEOUT,	c,	c
6	nextWorkerJob,	FIRST_JOB START_JOB,	w,	c

Σχήμα 3.5: Τμήμα προδιαγραφής μηνυμάτων συναλλαγών

Οι γραμμές 1 – 2 στο Σχήμα 3.5 δείχνουν ενέργεια που καλείται σε μεταβάσεις δύο διαφορετικών ρόλων. Στη γραμμή 3, λόγω της ενέργειας sendInitLog που εκτελεί ο συντονιστής της συναλλαγής (βλέπε Σχήμα 3.4 γραμμή 1), το συστατικό “lgc” απαντά με το μήνυμα INIT\_LOGGED. Για τον έλεγχο της εγκυρότητας του μοντέλου χρειαζόμαστε μόνο τα μηνύματα που καταναλώνονται από τους ρόλους του μοντέλου οπότε δεν είναι αναγκαίο να οριστεί το μήνυμα που στέλνει ο συντονιστής στο συστατικό “lgc”. Η γραμμή 5 δείχνει

μία ενέργεια που ακυρώνει το προγραμματισμένο μήνυμα TIMEOUT. Τέλος, η γραμμή 6 ορίζει μία ενέργεια που στέλνει είτε το μήνυμα FIRST\_JOB είτε το μήνυμα START\_JOB.

Σε αυτό το σημείο έχουμε πλέον τον πλήρη ορισμό του μοντέλου συναλλαγών αλλά και δύο αιτίες μη αιτιοκρατικής συμπεριφοράς. Η πρώτη αιτία προκύπτει από τον ορισμό των ρόλων του μοντέλου, όπως περιγράφεται στην Ενότητα 3.4. Η δεύτερη εισαγωγή έλλειψης αιτιοκρατίας προκαλείται από ενέργειες που παράγουν μηνύματα από ένα σύνολο εναλλακτικών επιλογών (βλέπε Σχήμα 3.5, γραμμή 6). Πιο στοιχείο επιλέγεται τελικά από το σύνολο των διαθέσιμων μηνυμάτων εξαρτάται από πληροφορίες που είναι διαθέσιμες κατά την εκτέλεση όπως, όπως το ιστορικό εκτέλεσης των επιμέρους εργασιών της συναλλαγής, λόγου χάρη.

Η δεύτερη αιτία μη αιτιοκρατικής συμπεριφοράς είναι απαραίτητο να απαλειφθεί ώστε να μην παράγονται λάθος αποτελέσματα κατά τον έλεγχο του μοντέλου. Για παράδειγμα, η ενέργεια nextWorkerJob στέλνει το μήνυμα FIRST\_JOB μόνο εάν πρόκειται για την πρώτη εργασία της συναλλαγής. Σε κάθε άλλη εργασία, το μήνυμα που λαμβάνει ο μέτοχος είναι το START\_JOB. Ο ορισμός των μηνυμάτων των συναλλαγών δε παρέχει αυτή την πληροφορία, οπότε ο ελεγκτής μοντέλου οφείλει να διερευνήσει όλα τα πιθανά ενδεχόμενα σε κάθε μετάβαση όπου εκτελείται η παραπάνω ενέργεια. Το σενάριο εκτέλεσης όπου ο μέτοχος λαμβάνει το μήνυμα START\_JOB πριν τη λήψη του μηνύματος FIRST\_JOB είναι σωστό με βάση τις πληροφορίες που διαθέτει ο ελεγκτής μοντέλου παρότι δεν μπορεί ποτέ να συμβεί. Η προσθήκη επιπλέον περιορισμών στη δημιουργία μηνυμάτων με βάση το ιστορικό εκτέλεσης, αφαιρεί τη παραπάνω μη αιτιοκρατική συμπεριφορά και επιλύει το πρόβλημα. Οι περιορισμοί παρέχονται σε δηλωτική μορφή στη γλώσσα προγραμματισμού με την οποία κατασκευάσαμε τον ελεγκτή μοντέλων (βλέπε Σχήμα 3.6).

```

valid path (Event "SPAWN" "c" "cn")      = (not o has_event_name "SPAWN") path
valid path (Event "SPAWN" "cn" "cn")    = False
valid path (Event "FIRST_JOB" "c" "w")  = (not o has_role "w") path
valid path (Event START_JOB "c" "w")    = has_event (Event "FIRST_JOB" "c" "w") path

```

Σχήμα 3.6: Περιορισμοί στην παραγωγή μηνυμάτων συναλλαγών

Ο πρώτος περιορισμός στο Σχήμα 3.6 αποτρέπει τη δημιουργία του μηνύματος SPAWN εάν έχει δημιουργηθεί πρωτύτερα μήνυμα αυτού του τύπου. Ο δεύτερος περιορισμός δεν επιτρέπει στους διαχειριστές υποσυναλλαγών να ξεκινήσουν νέες υποσυναλλαγές. Παρότι το μοντέλο ενθυλακομένων συναλλαγών επιτρέπει απεριόριστο βάθος συναλλαγών, για τον έλεγχο των ιδιοτήτων του μοντέλου αρκεί να εξετάσουμε συναλλαγές με βάθος δύο (η συναλλαγή και μία υποσυναλλαγή). Αν παραλειφθεί ο περιορισμός, ο ελεγκτής μοντέλου οφείλει να εξετάσει όλα τα πιθανά βάθη των δέντρων και δε θα τερμάτιζε ποτέ. Το μήνυμα FIRST\_JOB στέλνεται μόνο εάν ο μέτοχος της συναλλαγής δεν έχει λάβει κανένα μήνυμα από το συντονιστή. Τέλος, το μήνυμα START\_JOB λαμβάνεται μόνο αν έχει προηγηθεί το μήνυμα FIRST\_JOB. Σε σύνολο 32 ενεργειών που παράγουν μηνύματα, το ενθυλακόμενο μοντέλο παρουσιάζει έξι περιπτώσεις μη αιτοκρατικής συμπεριφοράς.

### 3.6 Πλήρης διερεύνηση σεναρίων εκτέλεσης

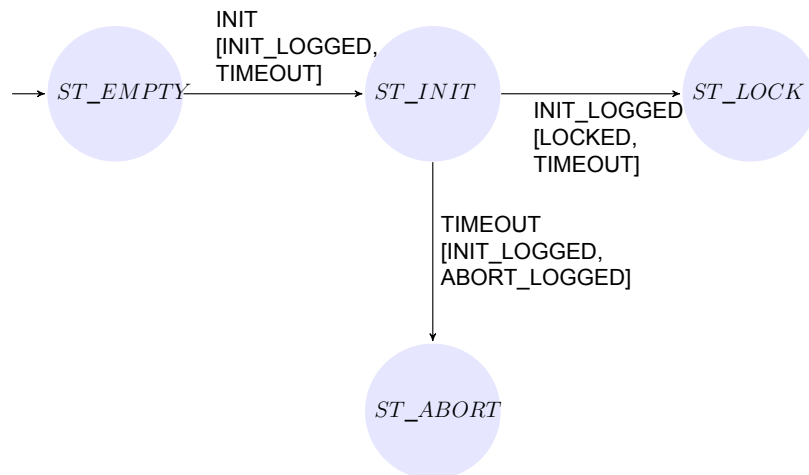
Η διερεύνηση των σεναρίων εκτέλεσης αφορά τον υπολογισμό του γραφήματος προ-σβασιμότητας του συγχρονισμένου γινομένου των αυτομάτων των ρόλων του μοντέλου [Bérard et al., 2001]. Η πρόσβαση των καταστάσεων επιτυγχάνεται:

- με το χειρισμό της λίστας των παραγόμενων μηνυμάτων που τη αποκαλούμε λίστα μελλοντικών μηνυμάτων (*fen*),
- τα μηνύματα που δεν έχουν σταλεί από ρόλο του μοντέλου αφαιρούνται από τη λίστα με βάση την σειρά προσέλευσης τους (FIFO) ενώ εξετάζονται όλοι οι πιθανοί συνδυασμοί σειράς προσέλευσης των μηνυμάτων που παράγονται από τους ρόλους,
- με το συγχρονισμό του καρτεσιανού γινομένου των αυτομάτων του μοντέλου με βάση τα μηνύματα συγχρονισμού.

Στο Σχήμα 3.7 βλέπουμε τις δύο αρχικές μεταβάσεις του αυτομάτου του συντονιστή συναλλαγών. Τα τόξα των μεταβάσεων αναγράφουν το μήνυμα που την προκάλεσε κα-

θώς και τα στοιχεία της λίστας μελλοντικών μηνυμάτων. Στην πρώτη μετάβαση, το μήνυμα TIMEOUT έχει παραχθεί από το συντονιστή ενώ το συστατικό "lgc" είναι ο δημιουργός του μηνύματος INIT\_LOGGED. Το γράφημα προσβασιμότητας συμπεριλαμβάνει όλους τους συνδυασμούς σειράς άφιξης των δύο μηνυμάτων:

- Η πρώτη πιθανή μετάβαση προκαλείται από την άφιξη του μηνύματος TIMEOUT. Το μήνυμα αφαιρείται από τη λίστα και προστίθεται το ABORT\_LOGGED που παράγεται από την τρέχουσα μετάβαση.
- Η άλλη μετάβαση προκαλείται από την άφιξη του μηνύματος INIT\_LOGGED.



Σχήμα 3.7: Γράφημα προσεγγισιμότητας του συγχρονισμένου γινομένου των αυτομάτων των ρόλων

Η μετάβαση προς την κατάσταση ST\_ABORT περιέχει δύο μελλοντικά μηνύματα που στέλνονται από το συστατικό του συστήματος υπεύθυνο για την καταγραφή πληροφοριών σε μόνιμο μέσο αποθήκευσης. Αφού δε πρόκειται για ρόλο του μοντέλου, τα δύο μηνύματα θα φτάσουν με σειρά FIFO, πρώτα θα φτάσει το μήνυμα INIT\_LOGGED και έπειτα το ABORT\_LOGGED.

Το βασικό μοντέλο συναλλαγών, όπως και το ενθυλακωμένο μοντέλο, περιλαμβάνει στον ορισμό του μεταβάσεις που δεν αλλάζουν την τρέχουσα κατάσταση και ούτε παράγουν

ούτε ακυρώνουν μηνύματα. Ο χρήστης ορίζει τις παραπάνω μεταβάσεις ώστε η σχέση μετάβασης του αυτομάτου του ρόλου να είναι ολικά καθορισμένη. Τις μεταβάσεις χωρίς επιπτώσεις στα παραγόμενα μηνύματα τις αποκαλούμε αδιάφορες και μπορούν να αφαιρεθούν από το μοντέλο πριν τον έλεγχο για την εγκυρότητα ιδιοτήτων αν:

1. Υπάρχουν και άλλες μεταβάσεις για την ίδια κατάσταση και το ίδιο μήνυμα. Σε μη αιτιοκρατικές μεταβάσεις παραλείπουμε εκείνες που δεν έχουν επιπτώσεις στη λίστα των μελλοντικών μηνυμάτων όταν υπάρχει τουλάχιστον μία μετάβαση που παράγει ή ακυρώνει μηνύματα. Αν όλες οι μεταβάσεις δεν επηρεάζουν τη λίστα, κρατάμε μόνο μία και παραλείπουμε όλες τις άλλες. Αν αγνοηθούν όλες η τελική κατάσταση του αυτομάτου θα ήταν πλέον απροσπέλαστη με αποτέλεσμα την εισαγωγή αδιεξόδου στο μοντέλο.
2. Το μήνυμα προκαλεί μόνο αδιάφορες μεταβάσεις σε όλα τα μονοπάτια που συνδέουν κάποια κατάσταση με την τελική κατάσταση του αυτομάτου. Όλες οι μεταβάσεις που προκαλούνται από το μήνυμα στις παραπάνω διαδρομές μπορούν να αγνοηθούν.

Η προεπεξεργασία του μοντέλου αποσκοπεί στον περιορισμό των μεταβάσεων και στη βελτίωση του χρόνου ελέγχου των ιδιοτήτων του μοντέλου. Η βελτίωση που τελικά επιτυγχάνεται εξαρτάται από το πλήθος των αδιάφορων μεταβάσεων που εκπληρώνουν τα παραπάνω κριτήρια.

Οι διαδρομές στο γράφημα προσβασιμότητας θεωρούνται κορεσμένες όταν ξεκινούν από την αρχική κατάσταση ενός ρόλου και τελειώνουν σε τελική κατάσταση του ίδιου ρόλου. Επιπλέον, στην τελευταία κατάσταση της διαδρομής η λίστα μελλοντικών μηνυμάτων πρέπει να είναι κενή. Τα κορεσμένα μονοπάτια αντιστοιχούν σε ολοκληρωμένες εκτελέσεις συναλλαγών. Η ορθότητα των ιδιοτήτων του μοντέλου ελέγχεται με βάση τις κορεσμένες διαδρομές που υπολογίζονται από τον αναδρομικό αλγόριθμο *path\_explore*.

### 3.6.1 Αλγόριθμος `path_explore`

Ο αναδρομικός αλγόριθμος υπολογίζει όλες τις κορεσμένες διαδρομές του μοντέλου και δέχεται τις παρακάτω παραμέτρους εισόδου:

- Δυάδες  $(q_r, fev_r)$  για κάθε ρόλο  $r$  όπου  $q_r$  είναι η τρέχουσα κατάσταση που εξερευνά ο αλγόριθμος ενώ η λίστα  $fev_r$  περιέχει τα μηνύματα που θα ληφθούν από το ρόλο. Στην πρώτη κλήση του αλγορίθμου, η  $q_r$  περιέχει την αρχική κατάσταση του αυτομάτου  $r$  ενώ η  $fev_r$  είναι κενή.
- Λίστα  $r_{trans}$  με τις μεταβάσεις όλων των ρόλων του μοντέλου όπως περιγράφονται στην Ενότητα 3.4.
- Λίστα  $ops$  με τις ενέργειες όπως περιγράφονται στην Ενότητα 3.5.
- Λίστα  $ignored$  που περιέχει τις καταστάσεις και τα μηνύματα που μπορούν να αγνοηθούν στην αντίστοιχη κατάσταση. Οι τιμές της λίστας προκύπτουν από την προεπεξεργασία σχετικά με τις αδιάφορες μεταβάσεις.
- Το μήνυμα  $ev$  που προκάλεσε τη μετάβαση. Το πρώτο μήνυμα που εξετάζει ο αλγόριθμος είναι το ειδικό μήνυμα INIT.
- Τη μη κορεσμένη διαδρομή  $path$  που επεκτείνει ο αλγόριθμος. Αρχικά η διαδρομή είναι κενή. Κάθε στοιχείο που προστίθεται έχει τη μορφή  $(q_{pre}, e, q_{post})$  και αντιστοιχεί στην μετάβαση από τη κατάσταση  $q_{pre}$  στη  $q_{post}$  που προκάλεσε το μήνυμα  $e$ .

Τα βήματα που ακολουθεί ο αλγόριθμος είναι:

1. Βρίσκουμε όλες τις μεταβάσεις στο  $r_{trans}$  που αποδέχονται το μήνυμα  $ev$ . Αν δε βρεθεί καμία τέτοια μετάβαση, ο ρόλος  $r$  έχει μερικώς ορισμένη σχέση μετάβασης και ο χρήστης ενημερώνεται για το λάθος του μοντέλου.
2. Για την αποτροπή κυκλικών διαδρομών, αφαιρούνται από τις διαθέσιμες μεταβάσεις που αποδέχονται το  $ev$  όσες απαντώνται στη διαδρομή  $path$ . Έστω  $ev_t$  το σύνολο



των μεταβάσεων που απομένουν. Αν υπάρχουν περισσότερες από μία μεταβάσεις, το μήνυμα  $ev$  προκαλεί μη αιτιοκρατική μετάβαση. Η αιτιοκρατία αποκαθίσταται στο επόμενο βήμα.

3. Για κάθε μετάβαση  $t$  από  $q_r$  σε  $q'_r$ , υπολογίζουμε το σύνολο των μηνυμάτων που παράγονται ή ακυρώνονται από τις ενέργειες που εκτελούνται στη μετάβαση.

(i) Έστω  $fev'_r$  η ενημερωμένη λίστα μελλοντικών μηνυμάτων που δίνεται από την εξίσωση:

$$fev'_r = \{x : x \in fev_r, -x \notin ev_t\} \oplus \{x : x \in ev_t, \neg cancel(x) \wedge P(x) \wedge (q_r, x) \notin ignored\}$$

όπου το  $P$  αναφέρεται σε συνθήκη που ορίζει ο χρήστης για την αποκατάσταση της αιτιοκρατικής επιλογής (βλέπε Σχήμα 3.6). Συμβολίζουμε με  $\oplus$  την παράθεση λιστών που διατηρεί τη σειρά των στοιχείων στην παράγωγο λίστα που είναι απαραίτητο για τη σειρά παράδοσης FIFO των μηνυμάτων που δε παράγονται από ρόλους.

(ii) Αν η κατάσταση  $q'_r$  είναι τελική, θέτουμε την αρχική κατάσταση του αυτομάτου στη  $q'_r$  και αφαιρούμε από την λίστα  $fev'_r$  τα μηνύματα με παραλήπτη το ρόλο.

4. Ενημερώνουμε τη μεταβλητή  $path'$  με την νέα μετάβαση ( $path' = path \oplus t$ ).

5. Για κάθε μήνυμα  $e \in fev'_r$  που παράγεται από ρόλο καθώς και για το πρώτο μήνυμα κάθε άλλης οντότητας που δεν είναι ρόλος, καλούμε τον αλγόριθμο `path_explore` με τιμές παραμέτρων:  $fev'_r \setminus \{e\}$  αντί για  $fev_r$ ,  $q'_r$  αντί για  $q_r$  και  $path'$  αντί για  $path$ .

6. Αν η λίστα  $fev'_r$  είναι κενή τότε:

(i) Η διαδρομή  $path$  είναι κορεσμένη αν για κάθε ρόλο  $k$  η λίστα  $fev_k$  είναι κενή.

(ii) Διαφορετικά, ο αλγόριθμος μεταβαίνει στο βήμα 5 αφού πρώτα θέσουμε την τιμή  $fev_k$  στη  $fev'_r$ .

Στο τρίτο βήμα του αλγορίθμου `path_explore` ερευνώνται μόνο οι μεταβάσεις που δεν έχουν ήδη επισκεφθεί σε προηγούμενη εκτέλεση του αλγορίθμου. Με αυτό τον τρόπο αποφεύγεται να διατρέξουμε πολλαπλά τους κύκλους του γραφήματος προσβασιμότητας. Η επιλογή μας αιτιολογείται από το γεγονός ότι κατά την εκτέλεση οι κύκλοι δεν ακολουθούνται για πάντα. Στο ρόλο του μετόχου για παράδειγμα, η ολοκλήρωση μίας εργασίας και η υποβολή νέας, σχηματίζουν κύκλο. Ωστόσο, οι συναλλαγές επαληθεύουν τις ιδιότητες του μοντέλου συναλλαγών ανεξάρτητα από το πλήθος εργασιών των μετόχων. Ο προτεινόμενος αλγόριθμος σπάει τους κύκλους του γραφήματος και τελικά εξερευνά το διευθυνόμενο άκυκλο γράφημα που προκύπτει. Μπορεί να δείχτεί ότι το γράφημα περιέχει πεπερασμένο πλήθος κορεσμένων διαδρομών οπότε η εξερεύνηση απαιτεί πεπερασμένο χρόνο.

Η έκρηξη του χώρου καταστάσεων είναι εγγενές πρόβλημα της μεθόδου ελέγχου μοντέλων. Σε περίπτωση εκδήλωσης του προβλήματος η μνήμη που απαιτείται για την αποθήκευση του γραφήματος προσβασιμότητας ή ο χρόνος υπολογισμού του γραφήματος, καθιστά ανέφικτη την εφαρμογή της μεθόδου. Το εργαλείο `ACID Model Checker` έχει υλοποιηθεί με τη συναρτησιακή γλώσσα προγραμματισμού `Haskell` [Peyton Jones, 2003]. Στη `Haskell` οι τιμές των παραμέτρων των συναρτήσεων υπολογίζονται όταν και εφόσον χρειαστούν σε άλλους υπολογισμούς (*lazy evaluation*). Η στρατηγική υπολογισμού της γλώσσας επιτρέπει την εύρεση και την αναφορά διαδρομών που παραβιάζουν τις εξεταζόμενες ιδιότητες, αλλά και άλλα προβλήματα του μοντέλου, προτού ολοκληρωθεί η κατασκευή του γραφήματος προσβασιμότητας. Παρότι δεν έχουμε αποκλείσει τη πιθανότητα εκδήλωσης έκρηξης του χώρου καταστάσεων, η δική μας υλοποίηση επιτρέπει τη μερική εξερεύνηση του μοντέλου. Το δεύτερο πλεονέκτημα είναι ότι τα κορεσμένα μονοπάτια διαγράφονται μετά τον έλεγχο για παραβιάσεις ιδιοτήτων με αποτέλεσμα τη μικρή κατανάλωση μνήμης. Επειδή το γράφημα δε περιέχει κύκλους, ο αλγόριθμος δεν μπορεί να διατρέξει πολλαπλά την ίδια διαδρομή και δεν υπάρχει λόγος να αποθηκεύονται οι διαδρομές που έχουν ήδη εξερευνηθεί.

Ο αλγόριθμος εξετάζει όλους τους πιθανούς συνδυασμούς της σειράς άφιξης των μηνυμάτων των ρόλων ακόμα και όταν η σειρά άφιξης δεν επηρεάζει την εγκυρότητα των

εξεταζόμενων μεθόδων. Σε μοντέλα με πολλούς ρόλους και μεταβάσεις, οι ανάγκες του αλγορίθμου σε υπολογιστική ισχύ και μνήμη είναι αποτρεπτικές. Η μείωση του πλήθους των καταστάσεων του μοντέλου με τη μέθοδο της μερικής διάταξης [Baier and Katoen, 2008] αποτελεί αποδεδειγμένα αποτελεσματική τεχνική για το περιορισμό των μεταβάσεων και του χρόνου ελέγχου του μοντέλου.

Για τον έλεγχο ιδιοτήτων του μοντέλου, αρκεί να εξετάσουμε το ελάχιστο αριθμό πραγμάτωσης των ρόλων που εκδηλώνουν όλες τις πιθανές επιδράσεις. Στο βασικό μοντέλο συναλλαγών υπάρχουν δύο οντότητες, ο συντονιστής και ο μέτοχος της συναλλαγής. Το ενθυλακόμενο μοντέλο έχει δύο επιπλέον ρόλους, το συντονιστή και το μέτοχο υποσυναλλαγών. Αρκεί να εξετάσουμε ένα αντιπρόσωπο από κάθε ρόλο. Οι ομοιότητες στη συμπεριφορά του συντονιστή συναλλαγής και του συντονιστή υποσυναλλαγών ή των μετόχων συναλλαγής και υποσυναλλαγών δεν μπορούν να αξιοποιηθούν στην τρέχουσα έκδοση του αλγορίθμου. Πιστεύουμε ότι η ομοιότητα των ρόλων μπορεί να χρησιμοποιηθεί αποτελεσματικά σε μελλοντικές βελτιώσεις του αλγορίθμου για τη περαιτέρω μείωση του χώρου καταστάσεων.

### **3.7 Επαλήθευση ιδιοτήτων μοντέλων κατανεμημένων συναλλαγών**

Το ACID Model Checker ανιχνεύει δομικά λάθη του μοντέλου καθώς και παραβιάσεις των αναμενόμενων ιδιοτήτων του. Τα δομικά λάθη περιλαμβάνουν καταστάσεις του μοντέλου που δεν είναι προσβάσιμες από την αρχική κατάσταση του αντίστοιχου αυτομάτου, μεταβάσεις που δεν μπορούν να ενεργοποιηθούν σε κανένα σενάριο εκτέλεσης και μερικά ορισμένη σχέση μετάβασης ενός ή παραπάνω αυτομάτων του μοντέλου. Η εγκυρότητα του μοντέλου αφορά χρονικές ιδιότητες που εκφράζουν την αναμενόμενη συμπεριφορά και εξέλιξη της κατάστασης του μοντέλου συναλλαγών στο χρόνο.

Απροσπέλαστες χαρακτηρίζονται οι καταστάσεις των αυτομάτων που δε μετέχουν σε

καμία από τις κορεσμένες διαδρομές του γραφήματος προσβασιμότητας. Οι απροσπέλαστες καταστάσεις υπολογίζονται από τη διαφορά του συνόλου των καταστάσεων των ρόλων από το σύνολο των καταστάσεων που απαντώνται σε μία τουλάχιστον κορεσμένη διαδρομή. Ο χρήστης παροτρύνεται να διορθώσει κατάλληλα το μοντέλο ώστε να μην υπάρχουν απροσπέλαστες καταστάσεις. Με παρόμοιο τρόπο υπολογίζονται και οι μεταβάσεις που δε μετέχουν σε καμία κορεσμένη διαδρομή.

Η σχέση μετάβασης του αυτομάτου είναι μερικώς ορισμένη όταν υπάρχει μήνυμα  $e$  που μπορεί να ληφθεί στην κατάσταση  $s$  αλλά δεν έχει οριστεί μετάβαση από τη  $s$  για το μήνυμα  $e$ . Ο αλγόριθμος αναφέρει την κατάσταση και το μήνυμα για τα οποία δεν έχει οριστεί η μετάβαση μόλις εντοπιστεί το σφάλμα στο μοντέλο χωρίς να περιμένει την ολοκλήρωση του ελέγχου. Στην ανάπτυξη του βασικού και του μοντέλου ενθυλακωμένων συναλλαγών διαπιστώσαμε ότι η ελλιπής σχέση μετάβασης είναι συχνή αιτία λαθών. Αν το λάθος δεν εντοπιστεί από τον έλεγχο μοντέλου κατά το σχεδιασμό, θα εκδηλωθεί στην εκτέλεση του προγράμματος. Η συμπεριφορά του αυτομάτου είναι απρόβλεπτη για απροσδιόριστες μεταβάσεις. Είναι εξαιρετικά δύσκολος ο εντοπισμός της αιτίας της απροσδόκητης συμπεριφοράς του αυτομάτου με βάση τη παρατηρούμενη συμπεριφορά κατά την εκτέλεση του προγράμματος.

Συνάρτηση	Περιγραφή
<code>has_event_name e path</code>	Μήνυμα με ίδιο τύπο όπως το $e$ έχει ληφθεί στη διαδρομή.
<code>has_event e path</code>	Το μήνυμα έχει ληφθεί στη διαδρομή.
<code>has_state s path</code>	Η κατάσταση $s$ υπάρχει στη διαδρομή.
<code>exists x path</code>	Ελέγχει εάν υπάρχει το $x$ στη διαδρομή. Το $x$ μπορεί να είναι μετάβαση, αρχική ή τελική κατάσταση ή μήνυμα.
<code>absent x path</code>	Η άρνηση της συνάρτησης <code>exists</code> .
<code>before x<sub>1</sub> x<sub>2</sub> path</code>	Το $x_1$ προηγείται του $x_2$ στη διαδρομή.
<code>after x<sub>1</sub> x<sub>2</sub> path</code>	Η άρνηση του <code>before</code> .

Πίνακας 3.1: Συναρτήσεις για τον ορισμό ιδιοτήτων ορθότητας

Στο ACID Model Checker, οι ιδιότητες του μοντέλου εκφράζονται από συναρτήσεις με τύπο  $path \rightarrow Bool$ , με πεδίο ορισμού τις διαδρομές του γραφήματος προσβασιμότητας και

σύνολο τιμών το σύνολο  $Bool$ . Οι ιδιότητες διατυπώνονται ως συνθέσεις των συναρτήσεων του Πίνακα 3.1 με λογικούς τελεστές (not, and, or).

Οι ιδιότητες ασφάλειας εκφράζουν τη βεβαιότητα ότι κάποιο ανεπιθύμητο γεγονός δεν μπορεί να συμβεί [Bérard et al., 2001]. Οι συναρτήσεις που εκφράζουν ιδιότητες ασφάλειας έχουν τύπο  $f : path \rightarrow Bool$  και επιστρέφουν  $True$  όταν η εξεταζόμενη ιδιότητα παραβιάζεται στη διαδρομή  $path$ .

Οι ιδιότητες προσεγγισιμότητας εκφράζουν την πεποίθηση ότι το σύστημα μπορεί να βρεθεί σε κάποια κατάσταση. Πρόκειται ουσιαστικά για την άρνηση ιδιοτήτων ασφάλειας και εκφράζονται από παρόμοιες συναρτήσεις που όμως επιστρέφουν την τιμή  $True$  για διαδρομές όπου ισχύει η εξεταζόμενη ιδιότητα.

Οι ιδιότητες βιωσιμότητας εκφράζουν την πεποίθηση ότι κάποιο γεγονός αναπόφευκτα θα συμβεί κατά την εκτέλεση του προγράμματος. Εκφράζονται από συναρτήσεις με τύπου  $f : path \rightarrow Bool$  που επιστρέφουν  $True$  αν το αναμενόμενο γεγονός δε βρεθεί στην εξεταζόμενη διαδρομή.

Έστω  $checked = \{p \in max\_paths : f(p) = True\}$  το σύνολο των κορεσμένων διαδρομών που επαληθεύονται από τη συνάρτηση  $f$ . Αν η συνάρτηση εκφράζει ιδιότητα ασφάλειας ή ιδιότητα βιωσιμότητας, τα στοιχεία του συνόλου είναι αντιπαραδείγματα που διαψεύδουν την εξεταζόμενη ιδιότητα. Αντίθετα, οι ιδιότητες προσεγγισιμότητας δεν ισχύουν αν το σύνολο είναι κενό.

### 3.7.1 Ιδιότητες μοντέλων κατανεμημένων συναλλαγών

Παρακάτω παρουσιάζουμε παραδείγματα ιδιοτήτων ασφαλείας, βιωσιμότητας και προσεγγισιμότητας για το βασικό και το μοντέλο ενθυλακομένων κατανεμημένων συναλλαγών.

Το Σχήμα 3.8 παρουσιάζει τις συναρτήσεις που εξετάζουν εάν οι μέτοχοι συναλλαγών εφαρμόζουν την απόφαση του συντονιστή (οριστικοποίηση αποτελεσμάτων ή ακύρωση εκτέλεσης). Η ιδιότητα πρέπει να επαληθεύεται και στα δύο μοντέλα συναλλαγών. Οι συναρτήσεις επιστρέφουν το σύνολο των εκτελέσεων όπου η απόφαση του συντονιστή δεν

εφαρμόζεται. Σε μοντέλα χωρίς σχεδιαστικά λάθη τα σύνολα αυτά πρέπει να είναι κενά.

```

1 both_commit p =
2   has_role      "w"                p &&
3   has_state     (State "ST_COMMIT_LOG" "c") p &&
4   has_not_event (Event "COMMIT" "c" "w")  p
5
6 both_abort p =
7   has_role      "w"                p &&
8   has_state     (State "ST_AKN_ABORT" "c") p &&
9   has_not_event (Event "ABORT" "c" "w")  p

```

Σχήμα 3.8: Ιδιότητα ασφαλείας: Οι μέτοχοι συναλλαγών εφαρμόζουν την απόφαση του συντονιστή

Η πρώτη συνάρτηση ορίζει τις συνθήκες όπου η εξεταζόμενη διαδρομή δεν έχει εφαρμόσει την απόφαση του συντονιστή για οριστικοποίηση των αποτελεσμάτων της συναλλαγής. Στη γραμμή 2 ελέγχουμε αν ο ρόλος του μετόχου συμμετέχει στην εξεταζόμενη διαδρομή. Υπενθυμίζουμε ότι ο ρόλος του μετόχου απουσιάζει από τοπικές συναλλαγές. Σε μη κατανεμημένες συναλλαγές η ιδιότητα δεν μπορεί να παραβιαστεί αφού δεν υπάρχει μέτοχος. Στη γραμμή 3 εξετάζουμε αν ο συντονιστής έχει αποφασίσει τη μονιμοποίηση της συναλλαγής έχοντας καταγράψει τη πρόθεσή του στο ιστορικό εκτέλεσης. Χωρίς την καταγραφή, η συναλλαγή είτε δεν έχει ολοκληρωθεί είτε πρόκειται να ακυρωθεί. Τέλος, στη γραμμή 4 εξετάζουμε αν ο μέτοχος έλαβε την απόφαση οριστικοποίησης των αποτελεσμάτων. Συνοψίζοντας, η συνάρτηση δηλώνει ότι εσφαλμένες είναι οι εκτελέσεις κατανεμημένων συναλλαγών όπου ο συντονιστής αποφάσισε την οριστικοποίηση των αποτελεσμάτων αλλά ο μέτοχος δεν έλαβε ποτέ την απόφαση.

Η συνάρτηση του Σχήματος 3.9 ελέγχει εάν ο μέτοχος οριστικοποιεί ή ακυρώνει τα αποτελέσματα της συναλλαγής. Με παρόμοιο τρόπο ελέγχουμε την ιδιότητα για το συντονιστή της συναλλαγής και, στο ενθυλακωμένο μοντέλο, για το συντονιστή και τον μέτοχο της υποσυναλλαγής.

Το Σχήμα 3.10 διατυπώνει την ιδιότητα ασφαλείας ότι οι υποσυναλλαγές ακυρωμέ-

```

1 worker_finished p =
2   has_role    "w"                p &&
3   has_event (Event "COMMIT" "c" "w") p ||
4   has_event (Event "ABORT" "c" "w") p

```

Σχήμα 3.9: Ιδιότητα ασφαλείας: Οι συναλλαγές ολοκληρώνονται (οριστικοποίηση ή ακύρωση αποτελεσμάτων)

```

1 top_nested_abort p =
2   has_role    "cn"                p &&
3   has_event (Event "ABORT" "c" "cn") p &&
4   has_not_state (State "ST_AKN_ABORT" "cn") p

```

Σχήμα 3.10: Ιδιότητα ασφαλείας: Οι υποσυναλλαγές ακυρώνονται αν ακυρωθεί η συναλλαγή

νων συναλλαγών ακυρώνονται επίσης [Chrysanthis and Ramamritham, 1990]. Παραβιάζουν την ιδιότητα οι εκτελέσεις όπου ο συντονιστής υποσυναλλαγών δεν εφαρμόζει την απόφαση του συντονιστή της συναλλαγής.

```

1 top_after_nested p =
2   has_role    "cn"                p &&
3   (
4     before_st (State "ST_AKN_COMMIT" "c" )
5               (State "ST_AKN_COMMIT" "cn") p ||
6     before_st (State "ST_AKN_COMMIT" "c" )
7               (State "ST_AKN_ABORT" "cn") p
8   )

```

Σχήμα 3.11: Ιδιότητα ασφαλείας: Η συναλλαγή μονιμοποιεί τα αποτελέσματα της αφού έχουν ολοκληρωθεί οι υποσυναλλαγές της

Η ιδιότητα του Σχήματος 3.11 ορίζει ότι η έκβαση της συναλλαγής υλοποιείται μετά την ολοκλήρωση των υποσυναλλαγών [Chrysanthis and Ramamritham, 1990]. Σε εκτελέσεις όπου παραβιάζεται η ιδιότητα, ο συντονιστής της υποσυναλλαγής αποφασίζει να οριστικοποίηση τα αποτελέσματα της υποσυναλλαγής χωρίς να περιμένει την απόφαση του συνο-

λικού συντονιστή.

```

1 independet_abort p =
2   has_role      "cn"                p &&
3   has_state     (State "ST_AKN_COMMIT" "c" ) p &&
4   has_state     (State "ST_AKN_ABORT"  "cn") p

```

Σχήμα 3.12: Ιδιότητα προσεγγισιμότητας: Οι υποσυναλλαγές ακυρώνονται ανεξάρτητα από την έκβαση της συναλλαγής

Στο Σχήμα 3.12 διατυπώνεται η ιδιότητα προσβασιμότητας ότι οι υποσυναλλαγές μπορούν να ακυρωθούν ανεξάρτητα από τη συναλλαγή [Chrysanthis and Ramamritham, 1990]. Μπορεί δηλαδή να ακυρωθεί η υποσυναλλαγή ακόμα και αν η συναλλαγή αποφασίσει την οριστικοποίηση των αποτελεσμάτων της. Η ιδιότητα ισχύει εάν το σύνολο των διαδρομών που επαληθεύουν τη συνάρτηση δεν είναι κενό.

```

1 coord_commit_abort p =
2   has_not_state (State "ST_AKN_COMMIT" "c") p &&
3   has_not_state (State "ST_AKN_ABORT"  "c") p

```

Σχήμα 3.13: Ιδιότητα βιωσιμότητας: Η συναλλαγή οριστικοποιεί ή ακυρώνει τα αποτελέσματά της

Η τελευταία ιδιότητα εκφράζει την πεποίθηση ότι οι συναλλαγές είτε οριστικοποιούν είτε ακυρώνουν τα αποτελέσματά τους σε όλες τις εκτελέσεις του μοντέλου. Οι συναλλαγές ολοκληρώνονται ακόμα και αν υπάρχουν αποτυχίες του περιβάλλοντος εκτέλεσης ή του δικτύου επικοινωνίας.

## 3.8 Σύνθεση κώδικα

Η μορφή του κώδικα που παρουσιάζουμε στο Σχήμα 3.14 είναι κατάλληλη για τον προσομοιωτή ACID Sim Tools. Ωστόσο, η ίδια μέθοδος μπορεί να εφαρμοστεί στη σύνθεση



κώδικα κάθε άλλου συστήματος επεξεργασίας συναλλαγών. Η πρόοδος των συναλλαγών παρακολουθείται από το παραγόμενο κώδικα με την καταγραφή της τρέχουσας κατάστασης των αυτομάτων του μοντέλου. Κατά τη μετάβαση των αυτομάτων των ρόλων σε νέα κατάσταση εκτελούνται οι ενέργειες που σχετίζονται με τη μετάβαση. Στην υλοποίηση του ACID Sim Tools, οι ενέργειες προσομοιώνουν υπηρεσίες οντοτήτων του μοντέλου όπως λόγω χάρη η καταχώριση εγγραφών στο ιστορικό και η διαχείριση κλειδιών διαμοιρασμένων αντικειμένων.

Οι συναλλαγές που φτάνουν στο σύστημα για εξυπηρέτηση σχετίζονται με αντικείμενα της κλάσης `CoordinatorTransactionFsm` που παρακολουθεί την εκτέλεσή της. Το αντικείμενο έχει πρόσβαση σε όλες τις πληροφορίες που αφορούν τη δομή της συναλλαγής όπως: (i) οι εργασίες της συναλλαγής, (ii) οι μέτοχοι που επιβεβαίωσαν την απόφαση του συντονιστή (iii) και οι μέτοχοι που ψήφισαν για την έκβαση της συναλλαγής. Αντίστοιχα, στους μετόχους οι συναλλαγές σχετίζονται με ένα αντικείμενο της κλάσης `TransactionFsm`.

```

for all role r in the model do
  if r is the event receiver then
    for all events ev triggering a transition of r do
      if ev is the received event then
        if the triggered transition is deterministic then
          set the current role state to the transition state
          invoke the operations associated with the transition
        else
          Let u be the result of the user–provided method
          for all state n accessible by ev from s do
            if n = u then
              set the current state of the role to n
              invoke the associated operations

```

Σχήμα 3.14: Πρότυπο σύνθεσης κώδικα για την εκτέλεση συναλλαγών

Σε μη αιτιοκρατικές μεταβάσεις, ο παραγόμενος κώδικας καλεί μεθόδους παρεχόμενες από το χρήστη ώστε να καθοριστεί η επόμενη κατάσταση του αυτόματου. Το όνομα της μεθόδου προκύπτει από τη συνένωση του ονόματος της τρέχουσας κατάστασης με το όνομα

του μηνύματος που προκαλεί τη μη αιτιοκρατική μετάβαση. Για παράδειγμα, η μέθοδος που καλείται από την κατάσταση `ST_VOTE` με μήνυμα `VOTED` ονομάζεται `resolveVotesVoted`. Η μέθοδος δέχεται ένα αντικείμενο της κλάσης `CoordinatorTransactionFsm` αν η μη αιτιοκρατική συμπεριφορά εκδηλώνεται στο συντονιστή ενώ όταν πρόκειται για το μέτοχο της συναλλαγής, το αντικείμενο είναι της κλάσης `TransactionFsm`. Η παρεχόμενη μέθοδος αξιοποιεί τις πληροφορίες που παρέχει το αντικείμενο παρακολούθησης ώστε να αποφασίσει την επόμενη κατάσταση του αυτόματου. Ο παραγόμενος κώδικας ελέγχει ότι η απόφαση της μεθόδου ανήκει στο σύνολο των εναλλακτικών καταστάσεων που ορίζει η μη αιτιοκρατική μετάβαση.

Το Σχήμα 3.14 δείχνει σε μορφή ψευδοκώδικα το πρότυπο κώδικα για τη σύνθεση της υλοποίησης της μεθόδου που χειρίζεται τα μηνύματα και τις μεταβάσεις του μοντέλου. Η μέθοδος χειρίζεται τα μηνύματα όλων των ρόλων του μοντέλου συναλλαγών και καλεί τις ενέργειες που σχετίζονται με τις εκτελούμενες μεταβάσεις. Φροντίζει επίσης για την αποκατάσταση της αιτιοκρατίας με τη βοήθεια των μεθόδων που παρέχει ο χρήστης.

## 3.9 Συμπεράσματα

Υπάρχει αυξανόμενο ενδιαφέρον για το σχεδιασμό και την ανάπτυξη μοντέλων συναλλαγών που χαλαρώνουν κάποιες από τις ACID ιδιότητες. Περιγράψαμε μια μέθοδο για την ανάπτυξη μοντέλων συναλλαγών που περιλαμβάνει ένα απλό και κατανοητό τρόπο για τον ορισμό μοντέλων, αλγοριθμικό έλεγχο των ιδιοτήτων του μοντέλου και την αυτόματη σύνθεση του κώδικα που επεξεργάζεται τα μηνύματα κατά την εκτέλεση της υλοποίησης. Δείξαμε την αναγκαιότητα της μη αιτιοκρατικής συμπεριφοράς του μοντέλου και δείξαμε πως να αποκαθίσταται η αιτιοκρατία στον έλεγχο του μοντέλου και στην παραγόμενη υλοποίηση.

Η προτεινόμενη προσέγγιση έχει εφαρμοστεί με επιτυχία στην επαλήθευση των ιδιοτήτων του πρωτοκόλλου 2PC [Mentis and Katsaros, 2009], του ενθυλακωμένου μοντέλου και στην υλοποίηση του προσομοιωτή ACID Sim Tools [Mentis et al., 2008], για την εκτίμηση

απόδοσης συστημάτων επεξεργασίας συναλλαγών [Mentis et al., 2009, Mentis et al., 2010].

Εξετάσαμε τρία είδη ιδιοτήτων: προσεγγισιμότητας, ασφάλειας και βιωσιμότητας. Στο μοντέλο ενθουλακομένων συναλλαγών όπου ο χώρος καταστάσεων είναι σημαντικά μεγαλύτερος, διαπιστώσαμε προβλήματα που δεν ήταν εμφανή στο βασικό μοντέλο. Επεκτείναμε την προτεινόμενη μέθοδο με περιορισμούς στη δημιουργία μηνυμάτων που βασίζονται στο ιστορικό της εξέλιξης της συναλλαγής. Η ενσωμάτωση προηγμένων τεχνικών περιορισμού του χώρου καταστάσεων του μοντέλου αναμένεται να βελτιώσει περαιτέρω την απόδοση της μεθόδου σε περίπλοκα μοντέλα συναλλαγών. Επίσης, η συντακτική υποστήριξη για τους περιορισμούς στη δημιουργία μηνυμάτων θα διευκολύνει τους χρήστες στη διατύπωση τους.

## **Κεφάλαιο 4**

# **Εκτίμηση απόδοσης συστημάτων επεξεργασίας συναλλαγών μέσω προσομοίωσης**

Τα συστήματα επεξεργασίας ασύγχρονων μηνυμάτων περιλαμβάνουν συστατικά που επικοινωνούν μέσω της ανταλλαγής μηνυμάτων όπου η σειρά λήψης τους ενδέχεται να είναι διαφορετική από τη σειρά αποστολής. Παραδείγματα τέτοιων συστημάτων συμπεριλαμβάνουν: συστήματα επεξεργασίας συναλλαγών, κατανεμημένα συστήματα αρχείων και οι υπηρεσίες πρόσβασης δεδομένων. Η πολυπλοκότητα των ασύγχρονων συστημάτων οφείλεται κυρίως σε τρεις λόγους: (i) υπάρχουν πολλές δυνατές καταστάσεις των συστατικών του συστήματος κατά τη λήψη του μηνύματος (ii) απαιτείται ακριβής καθορισμός της συμπεριφοράς σε περίπτωση αποτυχιών (iii) πρέπει να είναι εφικτή η ανάνηψη από αποτυχίες πτώσης του συστήματος. Η προσομοίωση προσφέρεται ως μέθοδος για την εκτίμηση της απόδοσης συστημάτων επεξεργασίας ασύγχρονων μηνυμάτων. Ωστόσο, για την εγκυρότητα των αποτελεσμάτων θα πρέπει να διασφαλιστεί η πιστή αναπαράσταση του συστήματος από το μοντέλο προσομοίωσης. Οι ιδιότητες ορθότητας που αναμένεται να

ισχύουν στο πραγματικό σύστημα πρέπει να επαληθεύονται και στο μοντέλο. Στόχος μας είναι η κατανόηση των αιτιών που καθορίζουν την απόδοση και τη διαθεσιμότητα του συστήματος. Με την κατάλληλη στατιστική επεξεργασία των αποτελεσμάτων επιδιώκουμε να αποτιμήσουμε αν εκπληρώνονται οι στόχοι της απόδοσης και της διαθεσιμότητας. Αν δεν μπορούν να εκπληρωθούν, τα αποτελέσματα της ανάλυσης πρέπει να μας καθοδηγούν προς τις κατάλληλες αλλαγές του μοντέλου ώστε να είναι εφικτή η εκπλήρωσή τους.

Στο παρόν κεφάλαιο προτείνουμε μία διαδικασία προσομοίωσης που αποσκοπεί στην απλοποίηση της εκτίμησης της απόδοσης περίπλοκων μοντέλων συστημάτων επεξεργασίας ασύγχρονων μηνυμάτων. Τα βήματα της διαδικασίας περιλαμβάνουν την περιγραφή του μοντέλου με μη αιτιοκρατικά αυτόματα, την επικύρωση των ιδιοτήτων του μοντέλου, την υποβοηθούμενη παραγωγή μέρους της υλοποίησης, την επικύρωση των πειραματικών παραμέτρων και των υποθέσεων του μοντέλου καθώς και τη στατιστική ανάλυση των αποτελεσμάτων της προσομοίωσης για την αποτίμηση της επιτευξιμότητας των στόχων απόδοσης. Η προτεινόμενη διαδικασία έχει εφαρμοστεί με επιτυχία στην εκτίμηση της επίδρασης των πρωτοκόλλων συναλλαγών και των σχετικών παραμέτρων στην απόδοση και τη διαθεσιμότητα των συστημάτων επεξεργασίας κατανεμημένων συναλλαγών.

Για την ανάλυση και την επικύρωση των μοντέλων επεξεργασίας κατανεμημένων συναλλαγών αξιοποιήσαμε τα εργαλεία ACID Model Checker [Mentis and Katsaros, 2012a] και ACID Sim Tools [Mentis et al., 2008, Mentis and Katsaros, 2012b]. Βασικές λειτουργίες, απαραίτητες κατά την εκτέλεση της προσομοίωσης, παρέχονται από συνοδευτική βιβλιοθήκη ρουτινών. Σε συστήματα επεξεργασίας συναλλαγών, οι βασικές λειτουργίες περιλαμβάνουν την καταγραφή δεδομένων σε μόνιμο μέσο αποθήκευσης και την προσωρινή αποκλειστική χρήση διαμοιρασμένων πόρων που εξασφαλίζεται από προσωρινό κλείδωμα των δεδομένων. Οι μεταβάσεις των αυτομάτων και οι πιθανές τους εκτελέσεις ελέγχονται για παραβιάσεις των προσδοκώμενων ιδιοτήτων του μοντέλου. Μέρος της υλοποίησης που είναι υπεύθυνη για το χειρισμό των αλληλεπιδράσεων μεταξύ των οντοτήτων του μοντέλου και την εκτέλεση των μεθόδων της συνοδευτικής βιβλιοθήκης, παράγεται αυτόματα από τον ορισμό του μοντέλου. Στα πειράματα προσομοίωσης εξετάζουμε με το τυπικά ελεγ-

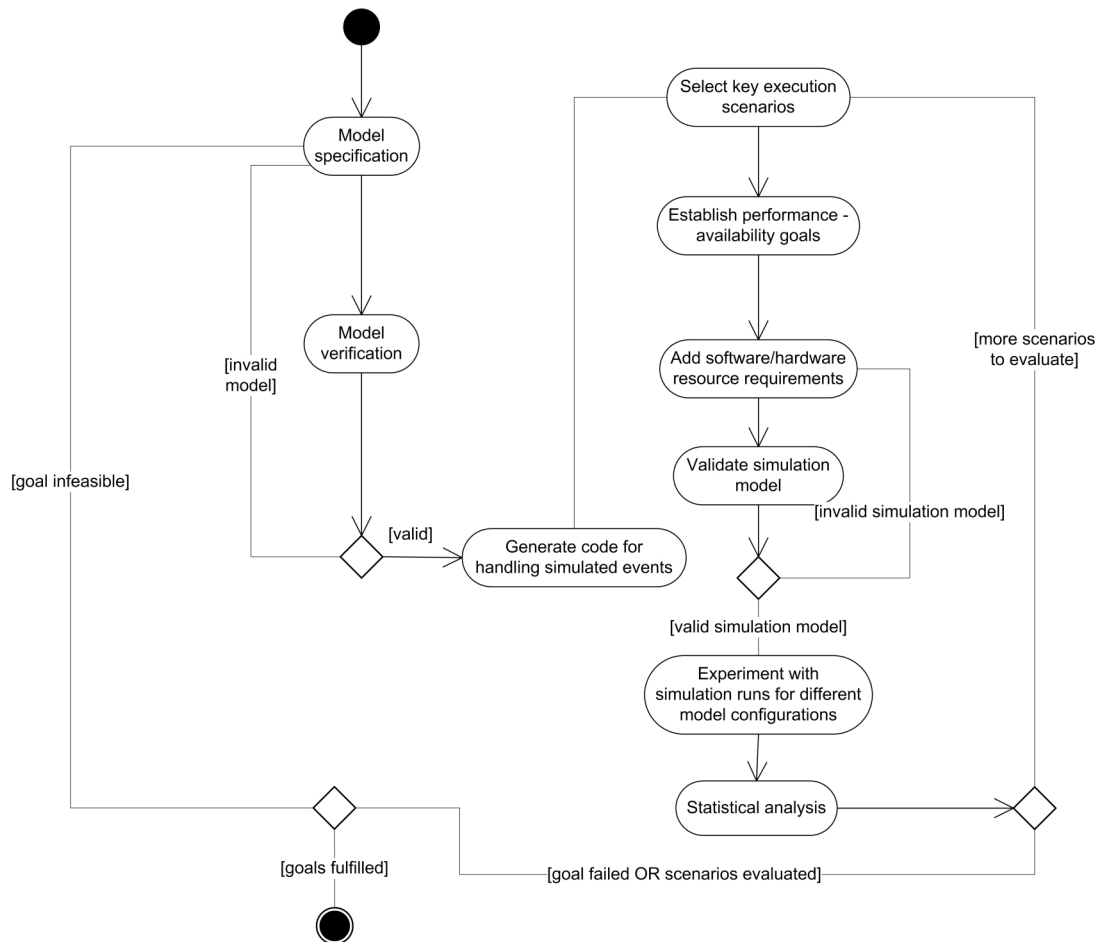
μένο προσομοιωτή βασικά ενδεχόμενα εκτέλεσης (λόγου χάρη σενάρια όπου εκτελούνται μόνο τοπικές συναλλαγές, ενδεχόμενα με υψηλές απαιτήσεις σε υπολογιστικούς πόρους) ώστε να διαπιστωθεί αν το μοντέλο μπορεί να εκπληρώσει τους στόχους απόδοσης και διαθεσιμότητας. Με κατάλληλη στατιστική επεξεργασία των πειραματικών αποτελεσμάτων εξετάζουμε τη δυνατότητα επιτυχίας των παραπάνω στόχων.

Τα πρώτα βήματα της προτεινόμενης διαδικασίας αξιοποιούνται στο [Mentis et al., 2010], για τον εντοπισμό σημείων ευαισθησίας και αντιστάθμισης. Εδώ εξετάζουμε μία ολοκληρωμένη διαδικασία προσομοίωσης και όχι τη στατιστική ανάλυση για την αποκάλυψη συσχέτισεων στις μετρικές των πειραματικών αποτελεσμάτων που περιγράφεται αναλυτικά στο Κεφάλαιο 5.

## 4.1 Διαδικασία ανάπτυξης μοντέλων προσομοίωσης

Στο Σχήμα 4.1 απεικονίζεται η διαδικασία προσομοίωσης για την αποτίμηση της απόδοσης και της διαθεσιμότητας κατανεμημένων συστημάτων με ασύγχρονη ανταλλαγή μηνυμάτων. Η διαδικασία περιλαμβάνει τα παρακάτω βήματα:

1. Ορισμός μοντέλου με μη αιτιοκρατικά αυτόματα που αντιστοιχούν σε οντότητες του συστήματος και περιγράφουν τις μεταξύ τους αλληλεπιδράσεις (βλέπε Ενότητα 3.4 και 3.5). Για παράδειγμα, το βασικό μοντέλο συναλλαγών δύο φάσεων 2PC, περιλαμβάνει δύο οντότητες, το συντονιστή και το μέτοχο κατανεμημένων συναλλαγών [Bernstein and Newcomer, 2009].
2. Έλεγχος για αποκλίσεις του μοντέλου από τις προδιαγραφές του συστήματος. Επιπλέον, το μοντέλο ελέγχεται εάν εκπληρώνει τις προσδοκώμενες ιδιότητες ορθότητας. Αν βρεθούν λάθη, προχωρούμε σε επαναδιατύπωση του μοντέλου ώστε να αντιμετωπιστούν τα σχεδιαστικά προβλήματα. Το βασικό μοντέλο συναλλαγών, για παράδειγμα, πρέπει να εκπληρώνει την ιδιότητα: “τελικά όλες οι συναλλαγές είτε οριστι-



Σχήμα 4.1: Αποτίμηση μέσω προσομοίωσης στόχων απόδοσης και διαθεσιμότητας σε συστήματα επεξεργασίας ασύγχρονων μηνυμάτων

κοποιούν είτε ακυρώνουν τα αποτελέσματά τους” [Mentis and Katsaros, 2009]. Αναλυτική περιγραφή του βήματος δίνεται στην Ενότητα 3.6.

3. Παραγωγή κώδικα για το χειρισμό των εισερχόμενων μηνυμάτων, των αλληλεπιδράσεων των οντοτήτων του μοντέλου και την κλήση των παρεχόμενων μεθόδων που παρέχει η συνοδευτική βιβλιοθήκη (βλέπε Ενότητα 3.8).
4. Επιλογή των βασικών ενδεχομένων εκτέλεσης που είτε θεωρούνται ως τυπικές περιπτώσεις χρήσης, είτε αναμένεται να εκτελούνται σπάνια αλλά έχουν σημαντικές επι-

πτώσεις στην απόδοση. Παραδείγματα τυπικών περιπτώσεων χρήσης από τις κατανεμημένες συναλλαγές είναι ενδεχόμενα που περιλαμβάνουν κατά κύριο λόγο τοπικές ή κυρίως κατανεμημένες συναλλαγές. Οι περιπτώσεις εμφάνισης αποτυχιών του εξυπηρετητή και η δημιουργία σημείου ελέγχου (περιοδική αφαίρεση από το ιστορικό καταχωρίσεων που δεν είναι πλέον χρήσιμες στην αποκατάσταση του εξυπηρετητή και των συναλλαγών), είναι ενδεχόμενα που πραγματοποιούνται σπάνια, με σημαντικές επιπτώσεις στη διαθεσιμότητα και την απόδοση.

5. Για κάθε ενδεχόμενο εκτέλεσης, διατυπώνονται οι στόχοι απόδοσης με βάση συγκεκριμένα ποσοτικά κριτήρια που πρέπει να εκπληρώνει η εξεταζόμενη εφαρμογή. Οι στόχοι ορίζονται ως όρια ελάχιστων ή μέγιστων τιμών των μετρικών όπως, λόγου χάρη, η ελάχιστη διαμεταγωγή συναλλαγών, ο μέγιστος χρόνος απόκρισης ή η ελάχιστη διαθεσιμότητα των εξυπηρετητών.
6. Ορισμός φόρτου εργασίας που περιλαμβάνει εργασίες με διαφορετικό ρυθμό άφιξης και με ποικίλες απαιτήσεις σε πόρους όπως:
  - (i) Απαιτήσεις λογισμικού όπως ο αναμενόμενος αριθμός εγγραφών ή ανάγνωσης δεδομένων από μόνιμο μέσο αποθήκευσης, περίοδος χρήσης της κεντρικής μονάδας επεξεργασίας και ο αριθμός των ανταλλασόμενων μηνυμάτων.
  - (ii) Υλικές απαιτήσεις όπως ο αριθμός των διαθέσιμων κεντρικών μονάδων επεξεργασίας και η ταχύτητα προσπέλασης των μόνιμων μέσων αποθήκευσης και του δικτύου.
7. Επικύρωση του μοντέλου με βάση:
  - (i) Την ικανότητα εντοπισμού των παραγόντων που διαδραματίζουν καθοριστικό ρόλο στην επίτευξη των στόχων απόδοσης.
  - (ii) Την προσέγγιση πραγματικών τιμών στις αριθμητικές τιμές των παραμέτρων του μοντέλου.

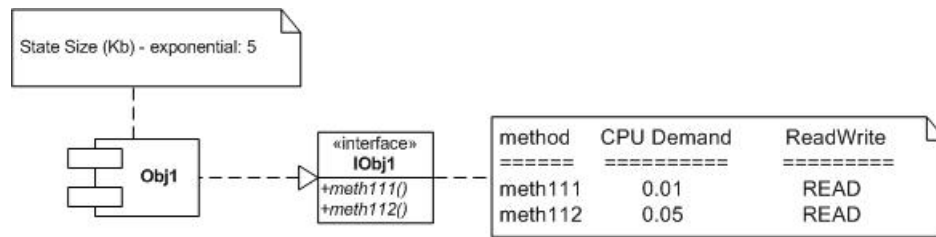


8. Εκτέλεση των προσομοιώσεων με διαφορετικούς συνδυασμούς τιμών των παραμέτρων του μοντέλου. Το σύνολο των συνδυασμών καθορίζουν ένα πειραματικό σχέδιο για την εξερεύνηση των επιπτώσεων των παραμέτρων στην απόδοση.
9. Στατιστική ανάλυση των πειραματικών αποτελεσμάτων για να εκτιμηθεί εάν εκπληρώνονται οι στόχοι απόδοσης. Η στατιστική ανάλυση αποκαλύπτει επίσης τους παράγοντες που έχουν σημαντικές επιδράσεις στην επίτευξη των στόχων. Όταν το μοντέλο αποτυγχάνει να εκπληρώσει τους στόχους, οι διαπιστωμένες επιδράσεις καθοδηγούν τον αναλυτή στο μερικό επαναπροσδιορισμό του μοντέλου. Στα συστήματα επεξεργασίας συναλλαγών, για παράδειγμα, αν η προσδοκώμενη διαμεταγωγή συναλλαγών δεν επιτευχθεί και βρεθεί ότι το πρωτόκολλο συναλλαγών αλληλεπιδρά με την επιλογή μέγιστου χρόνου ολοκλήρωσης συναλλαγών, η λύση θα μπορούσε να είναι η χρήση άλλου πρωτοκόλλου που μειώνει το χρόνο προσωρινής αποκλειστικής χρήσης διαμοιρασμένων αντικειμένων.

Τα τρία αρχικά βήματα της διαδικασίας περιγράφονται στο Κεφάλαιο 3. Στις επόμενες ενότητες δείχνουμε την εφαρμογή των υπολοίπων βημάτων με το προσομοιωτή ACID Sim Tools για την εκτίμηση της απόδοσης και της διαθεσιμότητας συστημάτων επεξεργασίας συναλλαγών.

## 4.2 Αφαιρέσεις προσομοίωσης κατανεμημένων συναλλαγών

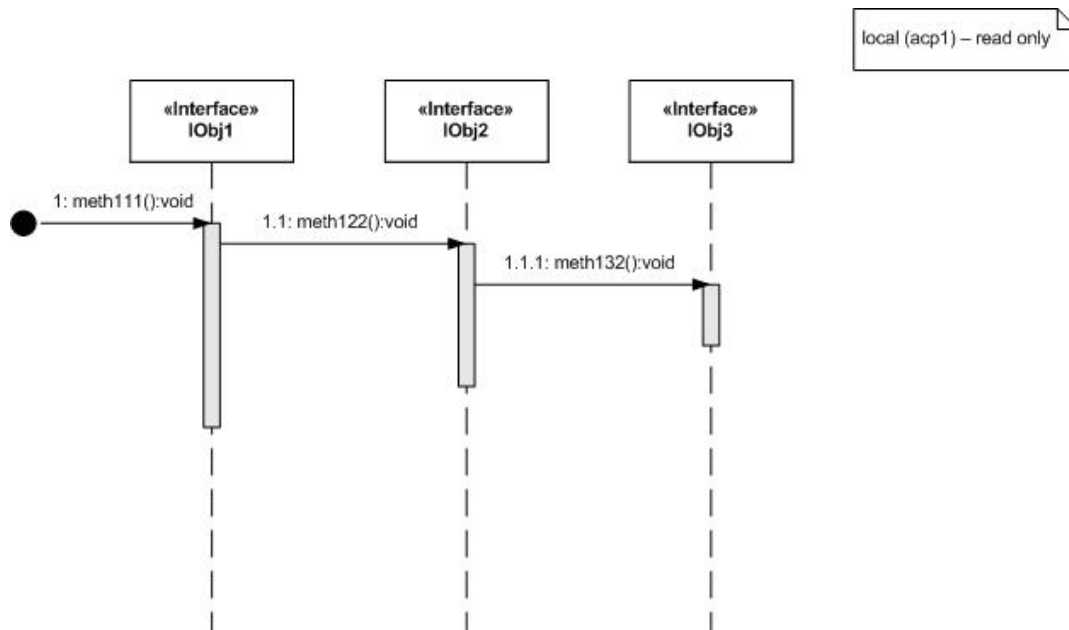
Οι αφαιρέσεις του μοντέλου προσομοίωσης κατανεμημένων συναλλαγών βασίζονται στις προδιαγραφές του OMG Core Object Model [Soley and Stone, 1995] και διασφαλίζουν την πιστή αναπαράσταση του αντικειμενοστραφή μοντέλου υπολογισμού. Οι χρήστες υποβάλλουν συναλλαγές για επεξεργασία στους εξυπηρετητές που διαχειρίζονται συλλογή από αντικείμενα και παρέχουν το περιβάλλον εκτέλεσης των αντικειμένων. Τα αντικείμενα



Σχήμα 4.2: UML διάγραμμα διεπαφής του αντικειμένου obj1

καλούν μεθόδους αντικειμένων που στεγάζονται στο ίδιο ή σε απομακρυσμένο εξυπηρετητή. Η εσωτερική κατάσταση των αντικειμένων μπορεί να μεταβληθεί μόνο από τις μεθόδους που παρέχει. Στο Σχήμα 4.2 βλέπουμε τη διεπαφή και τις μεθόδους ενός αντικειμένου καθώς και τις απαιτήσεις σε μνήμη για τη διατήρηση της εσωτερικής κατάστασης και τις ανάγκες σε υπολογιστική ισχύ για την εκτέλεση των μεθόδων. Το περιβάλλον εκτέλεσης του εξυπηρετητή προσφέρει την υπηρεσία αποθήκευσης και ανάκτησης της εσωτερικής κατάστασης του αντικειμένου σε μόνιμο μέσο αποθήκευσης. *Διαχρονικά* αποκαλούμε τα αντικείμενα που μπορούν να επαναφέρουν την εσωτερική κατάσταση από μόνιμο μέσο αποθήκευσης.

Η συναλλαγή είναι μια διαδικασία με πρόσβαση στην εσωτερική κατάσταση των αντικειμένων και έχει τη δυνατότητα να επιφέρει μεταβολές στην εσωτερική τους κατάσταση μέσω των μεθόδων τους [Martin and Ramamritham, 1997]. Το Σχήμα 4.3 απεικονίζει μία τοπική συναλλαγή και τα αντικείμενα που εμπλέκονται στην εκτέλεσή της. Αποτελείται από μεθόδους αντικειμένων που εκτελούνται σειριακά στους εξυπηρετητές όπου βρίσκονται τα αντικείμενα. Κάθε μέθοδος έχει συγκεκριμένες απαιτήσεις σε υπολογιστικούς πόρους που παρέχονται από μία ή παραπάνω κεντρικές μονάδες επεξεργασίας. Δυσλειτουργίες στην εκτέλεση των μεθόδων (exceptions) ή αδιεξόδων λόγω χρήσης διαμοιρασμένων αντικειμένων οδηγούν στην ακύρωση των εμπλεκόμενων συναλλαγών. Για να αποφεύγονται τα αδιέξοδα, οι συναλλαγές ακυρώνονται εάν δεν έχουν οριστικοποιήσει τα αποτελέσματά τους μετά την πάροδο ορισμένου χρόνου. *Τοπικές* αποκαλούνται οι συναλλαγές που καλούν μεθόδους αντικειμένων ενός μόνο εξυπηρετητή και *ανάγνωσης* είναι εκείνες που δε μεταβάλλουν την κατάσταση των αντικειμένων. *Αποκλεισμού* λέγονται οι συναλλαγές που μοιρά-



Σχήμα 4.3: Διάγραμμα ακολουθίας για τη συναλλαγή tr1

ζονται ένα ή παραπάνω αντικείμενα. Η εξάρτηση της λειτουργίας των συναλλαγών από τη διαθεσιμότητα των ίδιων αντικειμένων μπορεί να οδηγήσει σε αποκλεισμό από τη χρήση τους, αναστολή της λειτουργίας και τελικά στην ακύρωση των αποτελεσμάτων.

Ο διαχειριστής συναλλαγών παρέχει στις συναλλαγές που υποβάλλονται στον εξυπηρετητή για επεξεργασία τις ιδιότητες της Ατομικότητας, του Συγχρονισμού, της Απομόνωσης και της Οριστικότητας (ACID). Ο διαχειριστής υλοποιεί κάποιο πρωτόκολλο συντονισμού κατανεμημένων συναλλαγών όπως λόγου χάρη το πρωτόκολλο οριστικοποίησης δύο φάσεων (2PC) [Thanisch, 2000] ή κάποια παραλλαγή του.

Η απομόνωση των συναλλαγών διασφαλίζεται από τον έλεγχο πρόσβασης σε διαμοιρασμένα αντικείμενα. Στα πρωτόκολλα όπου τα αντικείμενα κλειδώνονται όταν απαιτείται προσωρινή αποκλειστική χρήση τους από κάποια συναλλαγή, αναστέλλεται η λειτουργία των υπόλοιπων συναλλαγών που χρειάζονται πρόσβαση στο ίδιο αντικείμενο μέχρι την άρση του κλειδώματος.

Ο περιορισμός των ενεργών νημάτων επεξεργασίας και κατά συνέπεια ο περιορισμός των συναλλαγών που εκτελούνται ταυτόχρονα στον εξυπηρετητή μπορεί να μετριάσει τη

συχνότητα ακύρωσης συναλλαγών λόγω αδιεξόδων. Οι συναλλαγές που βρίσκουν όλα τα νήματα επεξεργασίας κατειλημμένα, εισέρχονται σε ουρές αναμονής.

Κατά την κανονική λειτουργία των εξυπηρετητών καταγράφονται σε μόνιμο μέσο αποθήκευσης πληροφορίες απαραίτητες για την αποκατάσταση της λειτουργίας του εξυπηρετητή μετά την αποτυχία πτώσης και την αποκατάσταση της εσωτερικής κατάστασης αντικειμένων που μετέχουν σε ακυρωμένες συναλλαγές. Ο εξυπηρετητής και τα αντικείμενα που διαχειρίζεται δεν είναι διαθέσιμα κατά την ανάνηψη από αποτυχίες. Ο απαιτούμενος χρόνος για την ανάνηψη μπορεί να βελτιωθεί με τη περιοδική δημιουργία σημείου ελέγχου, την αφαίρεση δηλαδή των εγγραφών που δεν είναι πλέον χρήσιμες για την αποκατάσταση της λειτουργίας του εξυπηρετητή.

### 4.3 Ο προσομοιωτής ACID Sim Tools

Το ACID Sim Tools παρέχει εργαλεία για την κατασκευή, την επικύρωση και την εκτέλεση μοντέλων προσομοίωσης καταμετρημένων συναλλαγών. Υποθέτει την ύπαρξη αντικειμενοστραφή μοντέλου υπολογισμού, όπως διατυπώνεται από τις προδιαγραφές του OMG. Επεκτείνει τη λειτουργικότητα του περιβάλλοντος προσομοίωσης Objective Modular Network Testbed (OMNET++) [Varga and Hornig, 2008] και επιπλέον παρέχει:

- Γλώσσα για τον ορισμό μοντέλων και γραφική διεπαφή χρήστη για την εκτέλεση πειραμάτων και τον καθορισμό παραμέτρων του μοντέλου.
- Ιεραρχία κλάσεων που σε συνδυασμό με τον αυτόματα παραγόμενο κώδικα, επιτρέπει τον ορισμό, την επικύρωση και τη γρήγορη υλοποίηση προσομοιώσεων για νέα μοντέλα συναλλαγών.
- Ενσωματωμένες δυνατότητες αποσφαλμάτωσης μοντέλων προσομοίωσης και γραφική διεπαφή για τη διαδραστική εκτέλεση προσομοιώσεων και την εύκολη επισκόπηση του μοντέλου.

Δίνεται η δυνατότητα στο σχεδιαστή να συνδυάσει μοντέλα επεξεργασίας συναλλαγών με διαφορετικά μοντέλα διαχείρισης της πρόσβασης σε διαμοιρασμένα αντικείμενα και να πειραματιστεί με διαφορετικούς συνδυασμούς τιμών των παραμέτρων του σύνθετου μοντέλου. Έχουν ήδη υλοποιηθεί τα μοντέλα διαχείρισης συναλλαγών (i) βασικό 2PC PRN (καμία υπόθεση για την έκβαση της συναλλαγής) (ii) 2PC PRA (υποθέτει ακύρωση συναλλαγής) (iii) 2PC PRC (υποθέτει οριστικοποίηση συναλλαγής) (iv) και οι τρεις εκδοχές των παραπάνω πρωτοκόλλων για ενθυλακωμένες συναλλαγές. Έχουν επίσης υλοποιηθεί, η αυστηρή εκδοχή του κλειδώματος δύο φάσεων και η κατάταξη αντικειμένων με βάση χρονοσφραγίδες για τον έλεγχο της ταυτόχρονης πρόσβασης σε διαμοιρασμένα αντικείμενα. Το μοντέλο προσομοίωσης αποτελείται από αρθρώματα που υλοποιούν μέρος της λειτουργικότητας και επικοινωνούν μεταξύ τους με την ανταλλαγή μηνυμάτων. Στο ACID Sim Tools ορίζονται τα παρακάτω αρθρώματα:

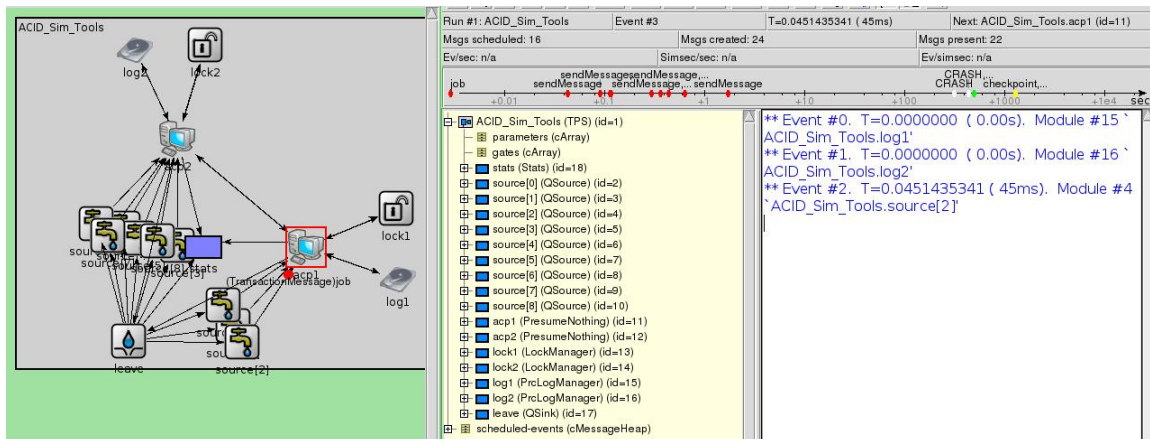
**QSource** Παράγει νέες αιτήσεις εξυπηρέτησης συναλλαγών με καθορισμένα χαρακτηριστικά που ονομάζουμε κλάσεις συναλλαγών. Τα χαρακτηριστικά που διαφοροποιούν τις κλάσεις συμπεριλαμβάνουν τα αντικείμενα της συναλλαγής, τις μεθόδους και τη σειρά εκτέλεσης τους.

**Atomic Commit Protocol (ACP)** Προσομοιώνει τις διεργασίες συντονισμού των συναλλαγών που επιτρέπουν τη λήψη συνολικής απόφασης για την έκβαση κατανεμημένης συναλλαγής. Τα αρθρώματα ACP μπορεί να υλοποιούν δύο διακριτούς ρόλους, το συντονιστή ή το μέτοχο σε συναλλαγή. Ο συντονιστής είναι υπεύθυνος για τη διεξαγωγή της ψηφοφορίας των μετόχων για να αποφασιστεί η έκβαση της συναλλαγής.

**Log Manager** Μοντελοποιεί το μέσο μόνιμης αποθήκευσης δεδομένων και διαχειρίζεται τις καταχωρίσεις των συναλλαγών που είναι απαραίτητες για την αποκατάσταση της λειτουργίας του εξυπηρετητή μετά από αποτυχία ή την επαναφορά της εσωτερικής κατάστασης των αντικειμένων ακυρωμένων συναλλαγών.

**Concurrency Control** Προσομοιώνει την πρόσβαση σε διαμοιρασμένα αντικείμενα και

παρέχει τις εγγυήσεις απομόνωσης [Atluri et al., 1997] της συναλλαγής. Όπως και στην εργασία [Chrysanthis et al., 1998], παραβλέπουμε τους υπολογιστικούς πόρους που καταναλώνονται στον έλεγχο της πρόσβασης λόγω της μεγάλης δυσκολίας ακριβής εκτίμησης των αναγκών σε υπολογιστικούς πόρους.



Σχήμα 4.4: Διάταξη των αρθρωμάτων προσομοίωσης (αριστερή πλευρά) και το διαδραστικό περιβάλλον προσομοίωσης του OMNET++ (δεξιά πλευρά)

Ο εξυπηρετητής συναλλαγών ορίζεται από τη σύνθεση των αρθρωμάτων ACP, Log Manager και Concurrency Control. Στο Σχήμα 4.4 βλέπουμε το γραφικό περιβάλλον εκτέλεσης προσομοιώσεων, τη διάταξη των αρθρωμάτων και τις μεταξύ τους συνδέσεις. Στο δεξί μέρος απεικονίζονται τα προγραμματισμένα μηνύματα που δεν έχουν ακόμα παραδοθεί στα κατάλληλα αρθρώματα καθώς και άλλες πληροφορίες χρήσιμες στην αποσφαλμάτωση μοντέλων προσομοίωσης.

Ο σχεδιαστής μπορεί να επιλέξει εναλλακτικές υλοποιήσεις των παραπάνω αρθρωμάτων, αρκεί να μην διαφοροποιείται το σύνολο των τύπων μηνυμάτων που αναγνωρίζει και στέλνει η εναλλακτική υλοποίηση.

Παρέχονται σειρά μετρικών που μπορεί να αξιοποιηθούν για τη μελέτη και την αντιμετώπιση προβλημάτων αντιστάθμισης, όπως η απόδοση και το κόστος της ανάνηψης από αποτυχία πτώσης. Υποστηρίζονται οι μετρικές:

- διαθεσιμότητα εξυπηρετητών
- διαμεταγωγή συναλλαγών (ποσοστό συναλλαγών που ολοκληρώνονται με επιτυχία) για τις ομάδες
  - (i) κατανεμημένων συναλλαγών
  - (ii) τοπικών συναλλαγών
  - (iii) και τις συναλλαγές που μοιράζονται ένα τουλάχιστον αντικείμενο
- μέσος χρόνος απόκρισης των παραπάνω ομάδων συναλλαγών
- μέσος χρόνος αποκλεισμού (το χρονικό διάστημα από τη δέσμευση του μετόχου της συναλλαγής με θετική ψήφο μέχρι τη λήψη της τελικής απόφασης από το συντονιστή) για τις παραπάνω ομάδες.

Ο προσομοιωτής υποστηρίζει τη μέθοδο των ανεξάρτητων επαναλήψεων για την ανάλυση των αποτελεσμάτων προσομοίωσης. Το πλήθος των προσομοιώσεων που είναι απαραίτητο για τη συλλογή αξιόπιστων αποτελεσμάτων καθορίζεται δυναμικά με βάση το απαιτούμενο ημιπλάτος του διαστήματος εμπιστοσύνης.

## 4.4 Εκτίμηση απόδοσης μοντέλου κατανεμημένων συναλλαγών

Σε αυτή την ενότητα παρουσιάζουμε την εφαρμογή της προτεινόμενης διαδικασίας προσομοίωσης στην εκτίμηση της απόδοσης και της διαθεσιμότητας εξυπηρετητών επεξεργασίας κατανεμημένων συναλλαγών. Από την εφαρμογή της διαδικασίας επιδιώκουμε να δείξουμε την αποτελεσματικότητά της στο εξεταζόμενο πρόβλημα αλλά και σε άλλα παρόμοια. Τέτοια συστήματα περιγράφονται από περίπλοκα μοντέλα προσομοίωσης όπου τα περισσότερα μηνύματα που ανταλλάσσονται μεταξύ των αρθρωμάτων της προσομοίωσης είναι

ασύγχρονα, σε αντιστοιχία με τα μηνύματα του πραγματικού συστήματος που προσομοιώνουν. Τα πρώτα τρία βήματα της διαδικασίας περιγράφονται αναλυτικά στο προηγούμενο κεφάλαιο. Εδώ επικεντρωνόμαστε στις αφαιρέσεις του μοντέλου προσομοίωσης και στην παρουσίαση της εφαρμογής των υπόλοιπων βημάτων στα πλαίσια του επιλεγμένου συνθετικού φόρτου εργασίας.

#### 4.4.1 Βασικά ενδεχόμενα εκτέλεσης

Στα βασικά ενδεχόμενα εκτέλεσης συμπεριλαμβάνονται φόρτοι εργασίας με διαφορετικό βαθμό κατανομής των συναλλαγών (τοπικές έως κατά κύριο λόγο κατανεμημένες συναλλαγές), συνδυασμούς από συναλλαγές που δεν μεταβάλλουν δεδομένα, από άλλες που μεταβάλλουν την εσωτερική κατάσταση των αντικειμένων, συναλλαγές αποκλεισμού που εντείνουν τον ανταγωνισμό για διαμοιρασμένα αντικείμενα και διαφορετικές συνθήκες ανταγωνισμού για κοινούς πόρους όπως πρόσβαση στα μόνιμα μέσα αποθήκευσης και υπολογιστικούς πόρους. Τα χαρακτηριστικά των συναλλαγών που μετέχουν στο συνθετικό φόρτο εργασίας παρουσιάζονται στον Πίνακα 4.1.

Ενδεχόμενα εκτέλεσης που συμβαίνουν σπάνια αλλά έχουν σημαντικές επιπτώσεις στην απόδοση και τη διαθεσιμότητα, συμπεριλαμβάνουν χαρακτηριστικά φόρτων εργασίας με συχνές αποτυχίες των εξυπηρετητών και συχνές λήψεις σημείων ελέγχου.

#### 4.4.2 Εκτίμηση απόδοσης και διαθεσιμότητας

Τυπικές απαιτήσεις από τα συστήματα επεξεργασίας κατανεμημένων συναλλαγών αφορούν τη διαμεταγωγή των συναλλαγών, το χρόνο απόκρισης και τη διαθεσιμότητα των εξυπηρετητών. Η διαμεταγωγή δίνεται από τον αριθμό των οριστικοποιημένων συναλλαγών στη μονάδα του χρόνου. Ο χρόνος απόκρισης ορίζεται ως ο μέσος χρόνος που παρέρχεται από τη στιγμή της υποβολή της συναλλαγής προς επεξεργασία μέχρι τη λήψη της απόφασης από το συντονιστή της έκβασης της συναλλαγής. Η διαθεσιμότητα υπολογίζεται από το



Κλάση	Καλούμενες μέθοδοι			Τοπική	Μεταβολή
tr1	meth111	meth122	meth132	Τοπική (acr1)	όχι
tr2	meth111	meth222	meth112	Κατανεμημένη	όχι
tr3	meth112	meth211	meth121	Κατανεμημένη	ναι
tr4	meth242	meth252	meth232	Τοπική (acr2)	όχι
tr5	meth242	meth142	meth242	Κατανεμημένη	όχι
tr6	meth242	meth141	meth251	Κατανεμημένη	ναι
tr7	meth262	meth272	meth282	Τοπική (acr2)	όχι
tr8	meth262	meth152	meth262	Κατανεμημένη	όχι
tr9	meth262	meth151	meth271	Κατανεμημένη	ναι

Πίνακας 4.1: Κλάσεις συναλλαγών, καλούμενες μέθοδοι και τρόπος πρόσβασης αντικειμένων

λόγο του χρόνου που ο εξυπηρετητής λειτουργεί κανονικά προς το συνολικό προσομοιωμένο χρόνο. Οι τιμές της διαθεσιμότητας εκφράζονται ως ποσοστό επί τοις εκατό.

Η διαμεταγωγή και ο χρόνος απόκρισης επηρεάζονται από τις επιλογές στην υλοποίηση των συστημάτων επεξεργασίας συναλλαγών και τις ρυθμίσεις των εξυπηρετητών, ενώ για την επίτευξη υψηλής διαθεσιμότητας απαιτείται γρήγορη αποκατάσταση της κανονικής λειτουργίας μετά από εκδήλωση αποτυχίας στον εξυπηρετητή. Με τη συσσώρευση μεγάλου αριθμού παρωχημένων εγγραφών στο ιστορικό των συναλλαγών αυξάνεται ο χρόνος που απαιτείται για την ανάνηψη του εξυπηρετητή και την αποκατάσταση των αντικειμένων που μετέχουν σε ακυρωμένες συναλλαγές. Μέχρι να ολοκληρωθεί η επαναφορά της εσωτερικής κατάστασης των αντικειμένων στην τελευταία γνωστή έγκυρη κατάσταση, τα αντικείμενα δεν είναι διαθέσιμα σε άλλες συναλλαγές με αποτέλεσμα τη μείωση της παρατηρούμενης διαμεταγωγής του εξυπηρετητή.

#### 4.4.3 Προσομοιωμένος συνθετικός φόρτος εργασίας

Στη μελέτη που παρουσιάζουμε εξετάζουμε βασικά ενδεχόμενα εκτέλεσης με συνθετικό φόρτο εργασίας που συμπεριλαμβάνει αντικείμενα τα οποία διαχειρίζονται δύο εξυπηρετητές. Στον Πίνακα 4.2 συνοψίζονται τα χαρακτηριστικά των αντικειμένων όπως ο εξυπηρε-

τητής που το διαχειρίζεται, οι απαιτήσεις σε μνήμη για τη διατήρηση της εσωτερικής κατάστασης, οι μέθοδοι και οι απαιτήσεις τους σε υπολογιστικούς πόρους. Επιπλέον, κάθε μέθοδος χαρακτηρίζεται με βάση το αν μεταβάλει την εσωτερική κατάσταση του αντικειμένου. Υποθέτουμε ότι το μέγεθος της εσωτερικής κατάστασης των αντικειμένων και οι απαιτήσεις σε υπολογιστικό χρόνο των μεθόδων έχουν εκθετική κατανομή με μέσο την αναγραφόμενη τιμή στο Πίνακα 4.2. Οι συναλλαγές καλούν τις μεθόδους με την αναγραφόμενη σειρά. Διακρίνουμε τις συναλλαγές σε κλάσεις ανάλογα με τις καλούμενες μεθόδους και τη σειρά εκτέλεσής τους. Η σειρά εκτέλεσης των μεθόδων καθορίζει και τη σειρά που επιχειρείται η απόκτηση προσωρινής αποκλειστικής χρήσης των αντίστοιχων αντικειμένων. Προτού κληθεί η μέθοδος, επιχειρείται κλειδωμά του αντικείμενο εφόσον είναι διαθέσιμο. Αφού ολοκληρωθεί η επεξεργασία της συναλλαγής, τα αντικείμενα ξαναγίνονται διαθέσιμα.

Οι υποθέσεις για τις ρυθμίσεις των εξυπηρετητών παρατίθενται στο Πίνακα 4.3. Θεωρούμε σταθερή την υστέρηση στη ανταλλαγή μηνυμάτων μεταξύ των εξυπηρετητών αφού οι εξυπηρετητές ανταλλάσσουν σύντομα μηνύματα και οι διακυμάνσεις στο χρόνο μετάδοσης είναι μηδαμινές.

Το βασικό μοντέλο επεξεργασίας συναλλαγών δύο φάσεων (2PC) και οι δύο παραλλαγές του, είναι πρωτόκολλα αποκλεισμού. Αν χαθεί προσωρινά η δυνατότητα επικοινωνίας μεταξύ των εξυπηρετητών, οι κατανεμημένες συναλλαγές είτε ακυρώνονται (αν ο εξυπηρετητής δεν έχει ψηφίσει για την έκβαση της συναλλαγής) είτε τα αντικείμενα που αφορούν τις αποκλεισμένες συναλλαγές παραμένουν κλειδωμένα και μη διαθέσιμα μέχρι την αποκατάσταση της επικοινωνίας. Ωστόσο, και τα τρία πρωτόκολλα μπορούν να αντιμετωπίσουν περιστασιακές αποτυχίες στην αποστολή μηνυμάτων.

Η περιστασιακή αποτυχία αποστολής μπορεί να προσομοιωθεί στο ACID Sim Tools θέτοντας μη μηδενική τιμή στην παράμετρο προσομοίωσης που αφορά την πιθανότητα αποτυχίας παράδοσης μηνυμάτων. Εμείς έχουμε θέσει μηδενική την παραπάνω πιθανότητα διότι ενδιαφερόμαστε κυρίως για την ανάνηψη από αποτυχίες του συστήματος και όχι για περιστασιακές αποτυχίες του δικτύου επικοινωνίας. Επιπλέον, και τα τρία πρωτόκολλα έχουν παρόμοια συμπεριφορά όσον αφορά την ανοχή σε μηνύματα που χάνονται.

Εξυπηρετητής	Αντικείμενο	Μέγεθος (Kb)	Μέθοδος	CPU	Πρόσβαση	
acr1	obj1	5	meth111	0.01	Ανάγνωση	
			meth112	0.05	Ανάγνωση	
			meth121	0.01	Μεταβολή	
	obj2	5	meth122	0.01	Ανάγνωση	
			meth123	0.01	Ανάγνωση	
			meth131	0.04	Μεταβολή	
	obj3	5	meth132	0.01	Ανάγνωση	
			meth141	0.01	Μεταβολή	
			meth142	0.01	Ανάγνωση	
	obj4	5	meth143	0.01	Ανάγνωση	
			meth151	0.01	Μεταβολή	
			meth152	0.01	Ανάγνωση	
	acr2	obj5	5	meth153	0.01	Ανάγνωση
				meth211	0.05	Μεταβολή
				meth212	0.05	Ανάγνωση
obj6		5	meth221	0.05	Μεταβολή	
			meth222	0.01	Ανάγνωση	
obj7		5	meth231	0.01	Ανάγνωση	
			meth232	0.01	Ανάγνωση	
obj8		5	meth241	0.05	Μεταβολή	
			meth242	0.05	Ανάγνωση	
obj9		5	meth251	0.05	Μεταβολή	
			meth252	0.01	Ανάγνωση	
obj10		5	meth261	0.01	Ανάγνωση	
			meth262	0.01	Ανάγνωση	
obj11	5	meth271	0.05	Μεταβολή		
		meth272	0.05	Ανάγνωση		
obj12	5	meth281	0.05	Μεταβολή		
		meth282	0.01	Ανάγνωση		
obj13	5					

Πίνακας 4.2: Απαιτήσεις σε υπολογιστικούς πόρους και μνήμη των διαχρονικών αντικειμένων

#### 4.4.4 Επικύρωση μοντέλου προσομοίωσης

Οι τιμές της χρονικής υστέρησης στην ανταλλαγή μηνυμάτων καθορίστηκαν από πειράματα σε πραγματικό δίκτυο. Η υποδομή του δικτύου και η απόσταση μεταξύ των σημείων επικοινωνίας επηρεάζουν καθοριστικά τις καταγεγραμμένες τιμές. Η μέση τιμή της υστέρησης βρέθηκε περίπου 60ms και συμφωνεί με σχετικά στατιστικά δεδομένα δημοσιευμένα από παρόχους δικτυακής πρόσβασης [Verizon, 2012]. Η ταχύτητα ανάγνωσης και εγγραφής δεδομένων σε μόνιμο μέσο αποθήκευσης επιλέχθηκε με βάση δημοσιευμένα χαρακτη-

ριστικά από εταιρίες κατασκευής σκληρών δίσκων. Τέλος, οι παράμετροι που αναφέρονται στους Πίνακες 4.2 και 4.1 αφορούν ειδικά χαρακτηριστικά των εφαρμογών που αξιοποιούν συναλλαγές και δεν χρειάζονται επικύρωση.

Υστέρηση ανταλλαγής μηνυμάτων		0.06 sec		
Εξυπηρετητής	Ταχύτητα δίσκου για		Μέσος χρόνος ανάνηψης	
	Ανάγνωση	Εγγραφή		
acp1	4.271e-05 s/Kb	51.252e-05 s/Kb	4 sec	
acp2	4.271e-05 s/Kb	51.252e-05 s/Kb	4 sec	

Πίνακας 4.3: Παράμετροι των εξυπηρετητών συναλλαγών

#### 4.4.5 Πειραματικός σχεδιασμός

Για τα βασικά ενδεχόμενα εκτέλεσης που περιγράφονται στις προηγούμενες ενότητες, εκτελέσαμε πλήρες πείραμα για όλους τους δυνατούς συνδυασμούς τιμών των πειραματικών παραμέτρων (βλέπε Πίνακα 4.4). Σε όλα τα πειράματα επιλέξαμε για την προσωρινή αποκλειστική χρήση διαμοιρασμένων αντικειμένων το πρωτόκολλο κλειδώματος δύο φάσεων (2PL) [Bernstein et al., 1987].

Το πλήρες πείραμα συμπεριλαμβάνει 486 εξεταζόμενες περιπτώσεις όπου ο προσομοιωμένος χρόνος είναι 55 ώρες και 30 λεπτά σε κάθε εκτέλεση του προσομοιωτή. Η συλλογή αποτελεσμάτων ξεκινά μετά την πάροδο ικανού χρόνου για την επίτευξη σταθερής κατάστασης. Ο χρόνος εκτέλεσης των πειραμάτων σε προσωπικό υπολογιστή με 1 GB διαθέσιμη μνήμη και ένα επεξεργαστή κυμάνθηκε από 2 έως 12 λεπτά ανάλογα με τις παραμέτρους του εκάστοτε πειράματος. Ο ρυθμός άφιξης των συναλλαγών καθορίζει κατά κύριο λόγο τον απαιτούμενο χρόνο για την εκτέλεση των πειραμάτων. Παρατηρήσαμε μικρή κατανάλωση μνήμης ( 5 – 50% της διαθέσιμης μνήμης) κατά την εκτέλεση των προσομοιώσεων και δεν παρουσιάστηκαν αστάθειες στη λειτουργία του προσομοιωτή ACID Sim

Πειραματικός Παράγοντας	Επίπεδο 1	Επίπεδο 2	Επίπεδο 3
Πρωτόκολλο οριστικοποίησης συναλλαγών (ACP)	Οριστικοποίηση δύο φάσεων – Υπόθεση έκβασης		
	Καμία (PRN)	Οριστικοποίηση (PRC)	Ακύρωση (PRA)
Επίπεδο πολυπρογραμματισμού (MPL)	2	3	4
Περίοδος δημιουργίας σημείου ελέγχου (CI) σε δευτερόλεπτα	500	1300	2100
Χρονικό όριο (TT)	0.9	1.1	1.3
Μέσος χρόνος άφιξης (MIT), εκθετική κατανομή	0.6	0.4	
Μέση περίοδος εμφάνισης αποτυχιών στους εξυπηρετητές (MITofSF), εκθετική κατανομή	18 λεπτά	5 ώρες 51 λεπτά	12 ώρες

Πίνακας 4.4: Πειραματικοί παράγοντες που επηρεάζουν την απόδοση και τη ταχύτητα ανάλυσης από αποτυχίες εξυπηρέτησης

Tools.

Ο *ομοιόμορφος πειραματικός σχεδιασμός* [Katsaros et al., 2007] προσφέρει μία εναλλακτική στρατηγική με σημαντικά μικρότερες απαιτήσεις σε υπολογιστικούς πόρους. Η προσομοίωση τερματίζεται όταν σημειωθεί διάστημα εμπιστοσύνης 95% στις μετρικές που μας ενδιαφέρουν, σύμφωνα με τη μέθοδο των ανεξάρτητων επαναλήψεων.

#### 4.4.6 Στατιστική ανάλυση

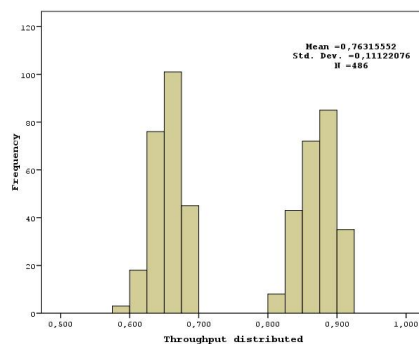
Η στατιστική ανάλυση των πειραματικών αποτελεσμάτων των προσομοιώσεων χρησιμοποιεί τη μέθοδο της πολυμεταβλητής ανάλυσης της διακύμανσης (MANOVA) [Harris, 2001]. Η μέθοδος αποσκοπεί στην ταυτόχρονη εξέταση πολλών πειραματικών αποτελεσμάτων και είναι επέκταση της παραγοντικής ανάλυσης διακύμανσης (ANOVA). Η MANOVA είναι ουσιαστικά πολυμεταβλητός στατιστικός έλεγχος που περιορίζει τη μεγέθυνση του ποσοστού λάθους που προκύπτει από τη διεξαγωγή πολλαπλών ελέγχων (ANOVA μιας μεταβλητής), όπου σε κάθε έλεγχο εξετάζει μία μεταβλητή (μετρική) [Katsaros et al., 2001]. Επίσης, η MANOVA εντοπίζει στα αποτελέσματα ομάδες μεταβλητών με αλληλεπιδράσεις αξιοποιώντας τις μεταξύ τους συσχετίσεις.

Διαπιστώνουμε ότι τα ιστογράμματα που προέκυψαν από τα πειραματικά αποτελέσματα από τα βασικά ενδεχόμενα εκτέλεσης των εξεταζόμενων μετρικών παρουσιάζουν

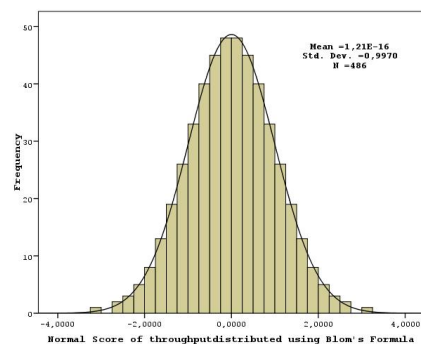
είτε μεγάλη ασυμμετρία είτε αφορούν πολυκόρυφη κατανομή (συνδυασμός πολλών κατανομών). Συμπεραίνουμε ότι τα δεδομένα δεν περιγράφονται από την κανονική κατανομή, όπως απαιτείται για την εφαρμογή της ανάλυσης διακύμανσης. Αξιοποιούμε το μετασχηματισμό Blom [Blom, 1958] για τη μετατροπή των αποτελεσμάτων στην απαιτούμενη μορφή. Ο μετασχηματισμός βασίζεται στις διαβαθμίσεις  $r_i$  των  $n$  τιμών και τη συσσωρευτική συνάρτηση της κανονικής κατανομής  $\Phi^{-1}(\cdot)$  (βλέπε Εξίσωση 4.1).

$$S_i = \Phi^{-1}\left(\frac{r_i - \frac{3}{8}}{n + \frac{1}{4}}\right) \quad (4.1)$$

Το ιστόγραμμα στο Σχήμα 4.5I δείχνει την κατανομή της συχνότητας με την οποία απαντώνται οι διαφορετικές τιμές της διαμεταγωγής κατανεμημένων συναλλαγών στα δεδομένα που προέκυψαν από τα πειράματα προσομοίωσης. Βλέπουμε από το Σχήμα 4.5II ότι η κατανομή των νέων μετασχηματισμένων μεταβλητών ταιριάζει πολύ καλά στην κανονική κατανομή με μέσο 0 και τυπική απόκλιση 1. Επιπλέον, οι επιδράσεις των πειραματικών παραγόντων διατηρούνται στα μετασχηματισμένα δεδομένα.



(I) Ιστόγραμμα αρχικών δεδομένων



(II) Μετασχηματισμένα δεδομένα

Σχήμα 4.5: Η κατανομή της συχνότητας εμφάνισης των τιμών της διαμεταγωγής κατανεμημένων συναλλαγών πριν και μετά το μετασχηματισμό Blom

Εξετάζουμε τις μετασχηματισμένες εξαρτημένες μεταβλητές αναφορικά με τους έξι πειραματικούς παράγοντες με τη στατιστική μέθοδο ανάλυσης MANOVA. Το μοντέλο ανάλυσης συμπεριλαμβάνει τους παράγοντες, τις 15 μη διατεταγμένες δυάδες των παραγόντων

και τις 20 μη διατεταγμένες τριάδες. Αλληλεπιδράσεις τεσσάρων ή παραπάνω παραγόντων δεν εξετάστηκαν λόγω της δυσκολίας που παρουσιάζει η ερμηνεία τους.

Διαπιστώνουμε από την τιμή της στατιστικής συνάρτησης ελέγχου F (F-test), ότι οι παρακάτω συνδυασμοί παραγόντων δεν έχουν σημαντικές επιπτώσεις στους στόχους απόδοσης και διαθεσιμότητας.

ACP * MPL * CI	ACP * MPL * MITofSF	ACP * CI * TT
ACP * CI * MIT	ACP * CI * MITofSF	MPL * CI * TT
CI * TT * MIT	ACP * MPL	

Πίνακας 4.5: Αλληλεπιδράσεις παραγόντων χωρίς σημαντικές επιπτώσεις στην απόδοση και τη διαθεσιμότητα εξυπηρέτησης

Οι πειραματικοί παράγοντες και οι συνδυασμοί παραγόντων που δεν αναγράφονται στον Πίνακα 4.5 έχουν στατιστικά σημαντικές επιπτώσεις στην απόδοση και τη διαθεσιμότητα των εξυπηρετητών. Σχετικά με την εγκυρότητα του στατιστικού μοντέλου, επιβεβαιώθηκε πως ταιριάζει καλά στα πειραματικά δεδομένα. Ο συντελεστής προσδιορισμού  $R^2$  δείχνει το ποσοστό της διακύμανσης που περιγράφει το στατιστικό μοντέλο. Η τιμή του συντελεστή προσεγγίζει τη μονάδα για κάθε εξαρτημένη μεταβλητή.

Για την αποτίμηση της επίπτωσης των παραγόντων στις εξαρτημένες μετρικές αξιοποιήσαμε μία σειρά από στατιστικούς ελέγχους που συνοδεύουν τη μέθοδο MANOVA. Οι έλεγχοι πραγματοποιούνται με το πρόγραμμα στατιστικής ανάλυσης SPSS [Field, 2009]. Στο Πίνακα 4.6 συνοψίζονται οι παράγοντες με στατιστικά σημαντική επίδραση στις εξαρτημένες μεταβλητές.

Το Σχήμα 4.6 δείχνει ενδεικτικά ραβδογράμματα της επίπτωσης των παραγόντων στις εξαρτημένες μεταβλητές. Το ύψος των ράβδων υποδεικνύει τη μέση τιμή των μετασχηματισμένων μεταβλητών αναφορικά με κάποιο παράγοντα.

Στα Σχήματα 4.6I και 4.6II βλέπουμε την επίπτωση της περιόδου δημιουργίας σημείου ελέγχου στη διαθεσιμότητα του εξυπηρετητή acp1 και στη διαμεταγωγή των κατανεμημένων συναλλαγών. Συχνές λήψεις σημείου ελέγχου βελτιώνουν τη διαθεσιμότητα με την αφαι-

Μετρική	ACP	MPL	CI	TT	MIT	MITofSF
Διαθεσιμότητα του εξυπηρετητή acp1	–	–	X	–	–	X
Διαθεσιμότητα του εξυπηρετητή acp2	–	–	X	–	–	X
Διαμεταγωγή τοπικών συναλλαγών	X	X	X	X	X	X
Διαμεταγωγή κατανεμημένων συναλλαγών	X	X	X	X	X	X
Μέσος χρόνος απόκρισης τοπικών συναλλαγών	X	X	X	X	X	X
Μέσος χρόνος απόκρισης κατανεμημένων συναλλαγών	X	X	X	X	X	X
Μέσος χρόνος αποκλεισμού	X	–	X	X	–	X

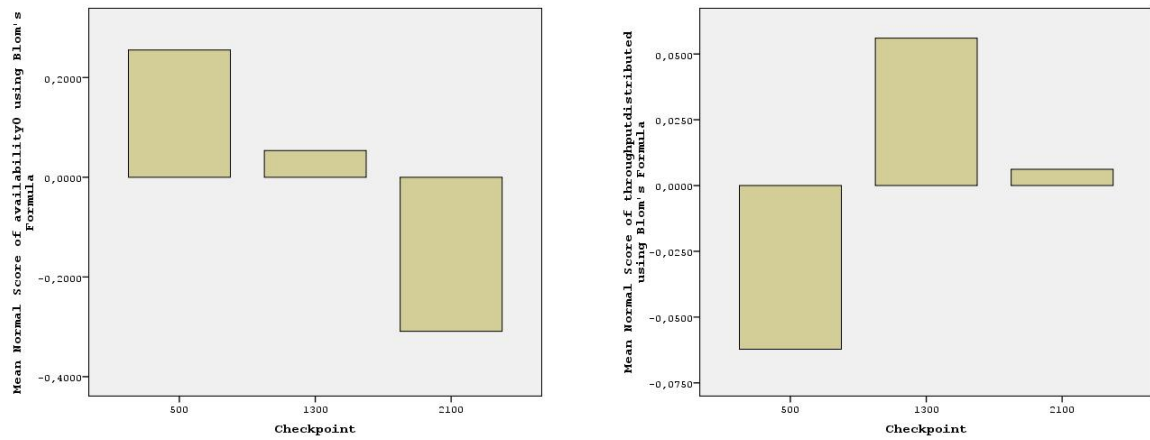
X = παράγοντας με στατιστικά σημαντική επίδραση  
 – = χωρίς επιδράσεις

Πίνακας 4.6: Παράγοντες με επιπτώσεις στη διαθεσιμότητα και την απόδοση συστημάτων επεξεργασίας κατανεμημένων συναλλαγών

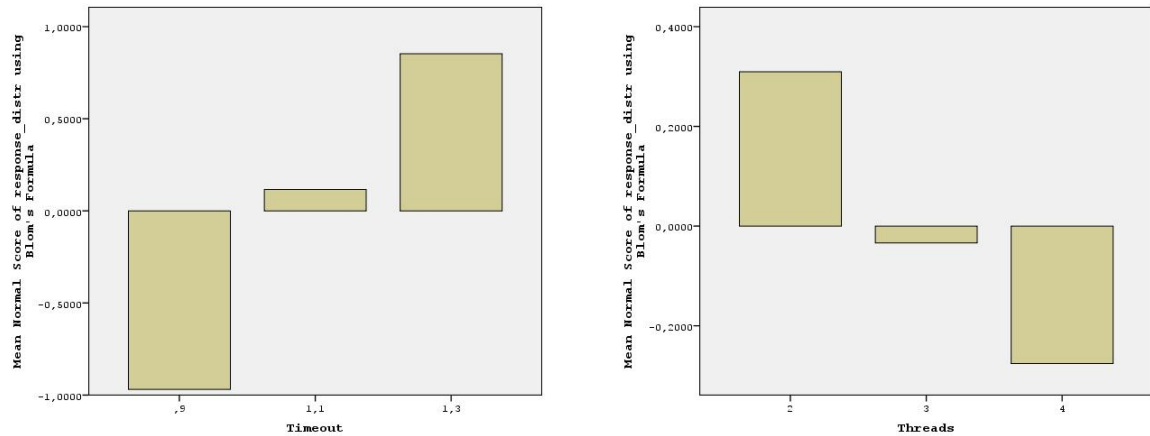
ρηση από το ιστορικό παρωχημένων εγγραφών και την επακόλουθη ταχύτερη ανάληψη του εξυπηρετητή. Από την άλλη, κατά τη δημιουργία σημείου ελέγχου, ο ανταγωνισμός για πρόσβαση στο μόνιμο μέσο αποθήκευσης αυξάνει τον απαραίτητο χρόνο για την επαναφορά της εσωτερικής κατάστασης των αντικειμένων ακυρωμένων συναλλαγών. Τελικά, όπως δείχνει το Σχήμα 4.6II, οι συχνές λήψεις σημείου ελέγχου μειώνουν τη διαμεταγωγή των συναλλαγών.

Η επίδραση του χρονικού περιθωρίου για την ολοκλήρωση συναλλαγών απεικονίζεται στο Σχήμα 4.6III. Με την αύξηση των χρονικών περιθωρίων, περισσότερες συναλλαγές κατορθώνουν να οριστικοποιήσουν τα αποτελέσματά τους. Άλλες μετρικές απόδοσης ωστόσο, μειώνονται σημαντικά (ο χρόνος απόκρισης για παράδειγμα). Από το Σχήμα 4.6IV, συμπεραίνουμε ότι υψηλές τιμές πολυπρογραμματισμού (αριθμός ενεργών νημάτων επεξεργασίας) βελτιώνει το χρόνο απόκρισης των κατανεμημένων συναλλαγών. Όταν υπάρχουν διαθέσιμοι υπολογιστικοί πόροι, οι συναλλαγές δε χρειάζεται να εισέρχονται σε ουρές αναμονής για να εξυπηρετηθούν οι υπολογιστικές τους απαιτήσεις.





(I) Διαθεσιμότητα και συχνότητα δημιουργίας ση- (II) Διαμεταγωγή και συχνότητα δημιουργίας μείου ελέγχου σημείου ελέγχου

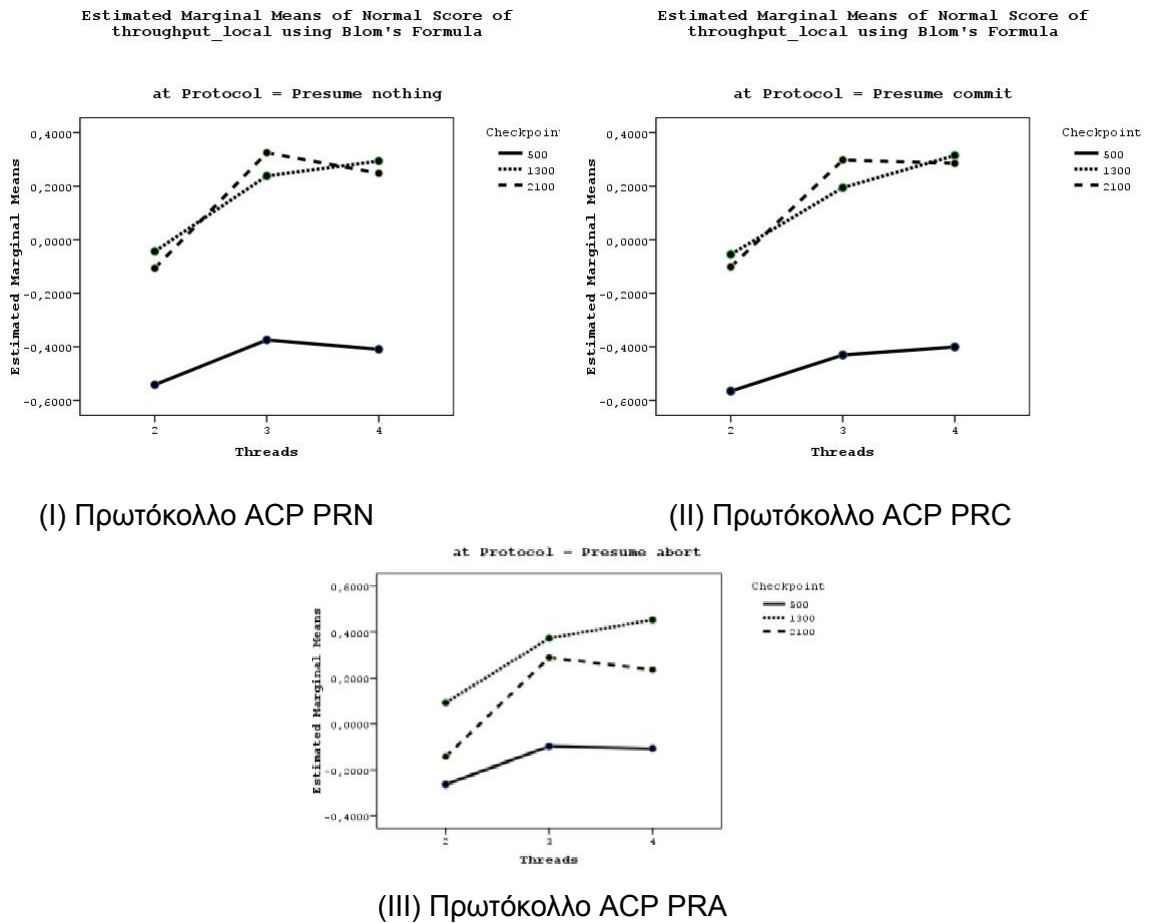


(III) Χρόνος απόκρισης και όριο ολοκλήρωσης συ- (IV) Χρόνος απόκρισης και αριθμός νημάτων επεξεργασίας

Σχήμα 4.6: Ραβδογράμματα των βασικών παραγόντων και οι επιδράσεις τους στην απόδοση και τη διαθεσιμότητα εξυπηρέτησης

Τα διαγράμματα των Σχημάτων 4.7 και 4.8 είναι κατάλληλα για την απεικόνιση των αλληλεπιδράσεων μεταξύ των παραγόντων. Αποτελούν χρήσιμο εργαλείο για την αποτίμηση των επιδράσεων διαφορετικών συνδυασμών πρωτοκόλλων και παραμέτρων των πρωτοκόλλων στους στόχους απόδοσης.

Το Σχήμα 4.7 αποκαλύπτει ενδιαφέρουσες αλληλεπιδράσεις μεταξύ των παραγόντων

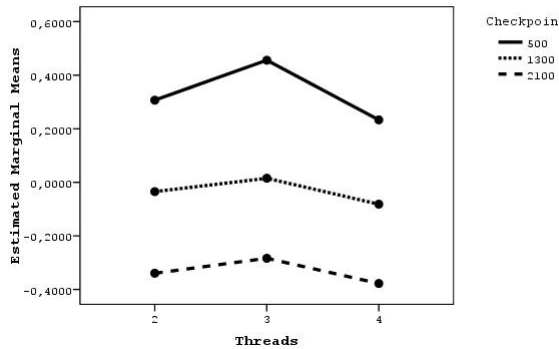


Σχήμα 4.7: Επίδραση πρωτοκόλλου συναλλαγών, συχνότητας λήψης σημείων ελέγχου και αριθμού νημάτων επεξεργασίας στη διαμεταγωγή τοπικών συναλλαγών

ACP, CI και MPL και τη συνδυασμένη επίδραση των παραγόντων στη διαμεταγωγή των τοπικών συναλλαγών. Παρατηρούμε σημαντική βελτίωση της διαμεταγωγής όταν η περίοδος δημιουργίας σημείου ελέγχου είναι 1300 δευτερόλεπτα και το πρωτόκολλο συντονισμού κατανεμημένων συναλλαγών είναι το ACP PRA. Υπάρχουν επίσης διαφορές στο ρυθμό διακύμανσης στα διαφορετικά επίπεδα των τιμών του CI όταν υπάρχουν περισσότερα ενεργά νήματα επεξεργασίας στον εξυπηρετητή. Τελικά, η καλύτερη απόδοση αναφορικά με τη διαμεταγωγή τοπικών συναλλαγών σημειώνεται για το πρωτόκολλο 2PC PRA, με τέσσερα ενεργά νήματα επεξεργασίας και δημιουργία σημείου ελέγχου κάθε 1300 δευτερόλεπτα.

Estimated Marginal Means of Normal Score of availability using Blom's Formula

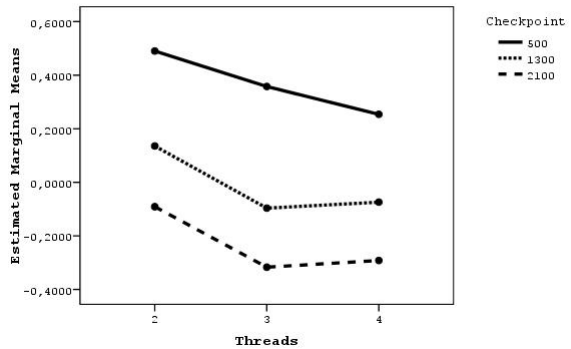
at Protocol = Presume nothing



(I) Πρωτόκολλο ACP PRN

Estimated Marginal Means of Normal Score of availability using Blom's Formula

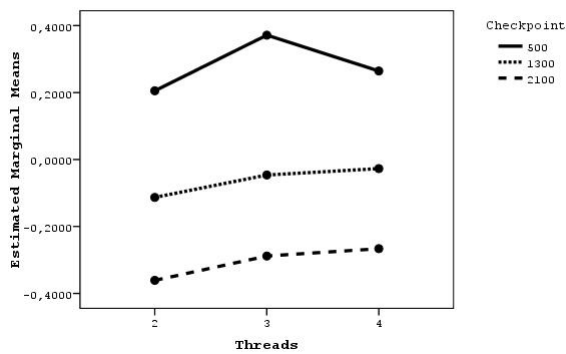
at Protocol = Presume commit



(II) Πρωτόκολλο ACP PRC

Estimated Marginal Means of Normal Score of availability using Blom's Formula

at Protocol = Presume abort



(III) Πρωτόκολλο ACP PRA

Σχήμα 4.8: Η επίδραση στη διαθεσιμότητα του εξυπηρετητή acp2 του πρωτοκόλλου συναλλαγών, συχνότητας λήψης σημείων ελέγχου και αριθμού νημάτων επεξεργασίας

Το Σχήμα 4.8 απεικονίζει της επιδράσεις των παραγόντων CI, MPL και ACP στη διαθεσιμότητα του εξυπηρετητή acp2. Παρατηρούμε ότι η διαθεσιμότητα του εξυπηρετητή μειώνεται όταν έχουμε συχνές λήψης σημείου ελέγχου με παραπάνω από τρία ενεργά νήματα επεξεργασίας. Το πρωτόκολλο ACP PRA παρουσιάζει ομοιόμορφη συμπεριφορά σε όλες τις τιμές των ενεργών νημάτων επεξεργασίας.

## 4.5 Συμπεράσματα

Σε αυτό το κεφάλαιο περιγράψαμε μία διαδικασία που βασίζεται σε προσομοίωση για την ανάλυση της απόδοσης συστημάτων διαχείρισης ασύγχρονων μηνυμάτων. Η προτεινόμενη διαδικασία ασχολείται με το πρόβλημα της ορθής μοντελοποίησης του συστήματος που επιδιώκουμε να μελετήσουμε, τον έλεγχο του μοντέλου ως προς τις ιδιότητες που αναμένεται να ικανοποιεί, την αυτόματη παραγωγή εγγυημένα σωστής υλοποίησης της διαχείρισης εισερχόμενων μηνυμάτων, την επιλογή βασικών ενδεχόμενων εκτέλεσης για τη μελέτη των στόχων απόδοσης, την εγκυρότητα των υποθέσεων του μοντέλου, τις τιμές των παραμέτρων προσομοίωσης, και τέλος, τη στατιστική επεξεργασία των αποτελεσμάτων για την εξαγωγή συμπερασμάτων σχετικά με τους στόχους απόδοσης και τη διαθεσιμότητα του συστήματος.

Η προτεινόμενη διαδικασία εφαρμόστηκε με επιτυχία σε βασικά ενδεχόμενα εκτέλεσης με συνθετικό φόρτο εργασίας κατανεμημένων συναλλαγών. Οι κατανεμημένες συναλλαγές είναι αντιπροσωπευτικό παράδειγμα πολύπλοκου συστήματος με ασύγχρονη επικοινωνία. Αντιμετωπίζουμε την εγγενή πολυπλοκότητα αυτών των συστημάτων και την επακόλουθη δυσκολία αποτίμησης της απόδοσης μέσω επικυρωμένων μοντέλων προσομοίωσης σχετικά με την πιστή αναπαράσταση των πολύπλοκων επιδράσεων που υπάρχουν στο σύστημα και τη στατιστική ανάλυση που συνυπολογίζει τα δεδομένα από πολλά πειράματα προσομοίωσης.

## **Κεφάλαιο 5**

# **Παράγοντες με σύνθετες αλληλεπιδράσεις στην απόδοση κατανεμημένων συναλλαγών**

Η αρχιτεκτονική λογισμικού ορίζεται από τις δομές που αποτελούν συστατικά λογισμικού του συστήματος (components), τις ορατές ιδιότητες των συστατικών και τις μεταξύ τους σχέσεις [Bass et al., 2003]. Η αρχιτεκτονική οριοθετεί την αναμενόμενη ποιότητα λογισμικού ως προς κάποια χαρακτηριστικά όπως (i) η απόδοση (ii) η ανοχή και η ανάκαμψη από λάθη (iii) ο διαμοιρασμός πόρων (iv) και η ασφάλεια, αλλά δεν μπορεί να αποτελέσει τη βάση για ακριβείς προβλέψεις [Dobrica and Niemelä, 2002] μιας και η συμπεριφορά του συστήματος εξαρτάται και από άλλους παράγοντες που αφορούν επιλογές υλοποίησης. Ωστόσο, η ανάλυση της αρχιτεκτονικής λογισμικού παρέχει τη δυνατότητα εξαγωγής φερέγγυων συμπερασμάτων σχετικά με την επίδραση σχεδιαστικών επιλογών σε ποιοτικά χαρακτηριστικά, όπως η αξιοπιστία λόγου χάρη.

Στα αποτελέσματα της έρευνας στην ποιότητα λογισμικού συμπεριλαμβάνονται πολυάριθμες εργασίες που αφορούν μεθόδους ανάλυσης της ποιότητας λογισμικού όπως οι:

ATAM [Clements et al., 2001], SBAR [Bengtsson and Bosch, 1998], SAAM [Kazman, 1994] και HoPLAA [Oliumofin and Mistic, 2007]. Όλες οι παραπάνω μέθοδοι συνδυάζουν ευριστικές ποιοτικές αναλύσεις, όπως ερωτηματολόγια και περιπτώσεις χρήσης του λογισμικού, με ποσοτικές αναλύσεις ειδικά προσαρμοσμένες στις ανάγκες της εκάστοτε μετρικής. Ωστόσο, η μεταβολή σε ένα και μόνο αρχιτεκτονικό χαρακτηριστικό μπορεί να επηρεάσει πολλές μετρικές λόγω των μεταξύ τους αλληλεπιδράσεων [Klein and Weiss, 2009].

Σε αυτό το κεφάλαιο προτείνουμε μία ποσοτική μέθοδο ανάλυσης για την ανακάλυψη συσχετίσεων μεταξύ των ποιοτικών μετρικών της αρχιτεκτονικής. Οι ποιοτικές μετρικές επηρεάζονται από το συνδυασμένο αποτέλεσμα ομάδας αρχιτεκτονικών χαρακτηριστικών όπως, λόγου χάρη, ο βαθμός παραλληλισμού της επεξεργασίας δεδομένων και ο ανταγωνισμός για πόρους διαμοιρασμένους μεταξύ των διεργασιών. Λανθάνοντες παράγοντες, που αν και επηρεάζουν τα ποιοτικά χαρακτηριστικά δεν μπορούν να καταμετρηθούν άμεσα, προκαλούν τους παρατηρούμενους συσχετισμούς. Περίπλοκες αντισταθμίσεις μεταξύ αντίρροπων ποιοτικών απαιτήσεων μπορούν να διαχειριστούν αξιοποιώντας τους λανθάνοντες ποιοτικούς παράγοντες.

Η μέθοδος εφαρμόστηκε στην αποτίμηση μέσω προσομοίωσης, αρχιτεκτονικών κατανεμημένων συναλλαγών που περιγράφονται από το πρότυπο του Διαχειριστή Διεργασιών [Gorton, 2006]. Το πρότυπο συχνά χρησιμοποιείται για στην υλοποίηση επιχειρησιακών διεργασιών που υποβάλλουν αιτήματα σε ένα ή περισσότερους εξυπηρετητές. Ζητήματα σχετικά με την αποδόμηση της λειτουργικότητας, τη δέσμευση διαμοιρασμένων πόρων και την επικοινωνία μεταξύ των συστατικών του εξυπηρετητή, οριοθετούν ένα δύσκολο και ενδιαφέρον πρόβλημα σχεδίασης. Εργασίες που αφορούν εφαρμογές αρχιτεκτονικών κατανεμημένων συναλλαγών [Bass et al., 2003] αναφέρουν μεγάλες διακυμάνσεις στη διαθεσιμότητα του εξυπηρετητή, την απόδοση και τη δυνατότητα κλιμάκωσης του φόρτου εργασίας. Οι μετρικές ποιότητας που αναλύονται σε αυτό το κεφάλαιο αφορούν και άλλα αρχιτεκτονικά πρότυπα όπως τα Broker και Publish–Subscribe [Gorton, 2006].

Στο πρώτο στάδιο της προτεινόμενης μεθόδου διεξάγεται παραγοντική ανάλυση η οποία αποκαλύπτει ομάδες μετρικών με υψηλή συσχέτιση και τις αντικαθιστά με νέες ασυσχέτιστες

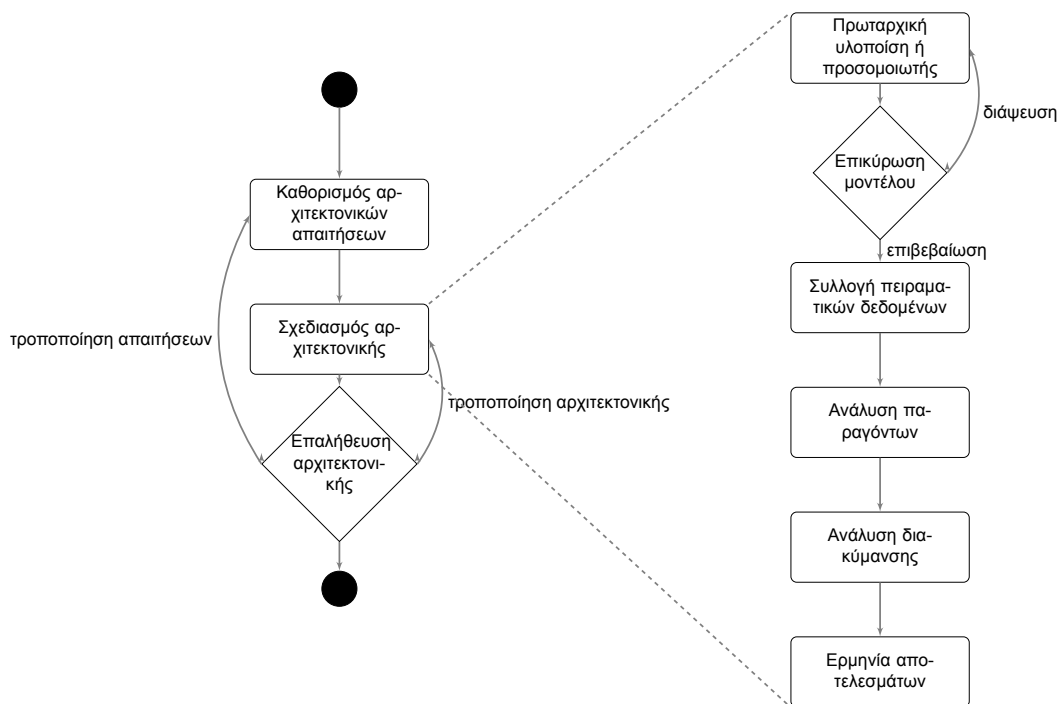
μεταβλητές που αντιστοιχούν στην επίδραση των λανθάνοντων ποιοτικών παραγόντων της αρχιτεκτονικής. Ανάμεσα στις ομαδοποιημένες μετρικές υπάρχει θετική ή αρνητική συσχέτιση. Στην ανάλυση που παρουσιάζουμε, η παραγοντική ανάλυση εφαρμόστηκε σε δεδομένα που προέκυψαν από πολλές και ανεξάρτητες εκτελέσεις πειραμάτων προσομοίωσης με το προσομοιωτή επεξεργασίας συναλλαγών ACID Sim Tools [Mentis et al., 2008].

Στο δεύτερο στάδιο της μεθόδου, εκτελούνται ξεχωριστές εκτιμήσεις διακύμανσης (ANOVA) για κάθε λανθάνων ποιοτικό χαρακτηριστικό που φανέρωσε το πρώτο στάδιο της μεθόδου.

Η προτεινόμενη μέθοδος σχετίζεται με συστηματικό τρόπο τη διακύμανση των τιμών των ποιοτικών μετρικών και τις επιλογές αντιστάθμισης, με τα αρχιτεκτονικά χαρακτηριστικά που τα προκαλούν. Τα δεδομένα της ανάλυσης προέρχονται από κατάλληλο περιβάλλον προσομοίωσης [Balsamo and Marzolla, 2003, Balsamo et al., 2004b]) ή από μία πρωτότυπη και ημιτελή υλοποίηση [Bardram et al., 2005, Becker et al., 2008, Vieira and Madeira, 2003]). Βασίζεται σε ελάχιστες υποθέσεις σχετικά με τις κατανομές που περιγράφουν τα δεδομένα και δεν εξαρτάται από το εξεταζόμενο ποιοτικό χαρακτηριστικό ή τα αρχιτεκτονικά χαρακτηριστικά που το επηρεάζουν. Έχουμε την πεποίθηση ότι η μέθοδος μπορεί να εφαρμοστεί σε κάθε αρχιτεκτονική ανεξάρτητα από τις αρχιτεκτονικές μορφές που την περιγράφουν, αρκεί να υπάρχουν αλληλεπιδράσεις μεταξύ των ποιοτικών μετρικών. Οι ελάχιστες προϋποθέσεις για την εφαρμογή της μεθόδου αιτιολογούν την πεποίθησή μας, αλλά η αποτελεσματικότητά σε άλλες αρχιτεκτονικές, πέρα των κατανεμημένων συναλλαγών, μένει να επιβεβαιωθεί σε μελλοντικές ερευνητικές εργασίες.

## 5.1 Διαδικασία εκτίμησης αρχιτεκτονικών συναλλαγών

Η διαδικασία για την αποτίμηση αρχιτεκτονικών λογισμικού περιλαμβάνει στάδια που συμβάλουν στο σχεδιασμό και την επικύρωση της ως προς τις λειτουργικές και ποιοτικές απαιτήσεις.



Σχήμα 5.1: Προτεινόμενη μέθοδος στα πλαίσια της γενικής διαδικασίας αποτίμησης αρχιτεκτονικών

Στο [Gorton, 2006] παρουσιάζεται μία διαδικασία αποτίμησης αρχιτεκτονικών που απεικονίζεται στο Σχήμα 5.1 και περιλαμβάνει τρία στάδια: (i) τον προσδιορισμό των αρχιτεκτονικών απαιτήσεων (ii) το σχεδιασμό της αρχιτεκτονικής (iii) και την επικύρωσή της.

Τα τρία στάδια της διαδικασίας εκτελούνται επαναληπτικά, αφού η επικύρωση της αρχιτεκτονικής ενδέχεται να μην ικανοποιήσει τις προσδοκίες, οπότε απαιτείται αλλαγή του σχεδιασμού της αρχιτεκτονικής ή επαναπροσδιορισμός των απαιτήσεων.

Η προτεινόμενη μέθοδος ανάλυσης εφαρμόζεται στο τρίτο στάδιο της διαδικασίας αποτίμησης. Τα βήματα της μεθόδου απεικονίζονται στο δεξί τμήμα του Σχήματος 5.1. Άλλα σημαντικά ζητήματα σχεδιασμού όπως η εφαρμογή αρχιτεκτονικών αρχέτυπων ή στρατηγικές για τη βελτίωση του αρχιτεκτονικού σχεδιασμού [Williams and Carver, 2010], δεν αντιμετωπίζονται από την προτεινόμενη μέθοδο. Η αντιμετώπιση των παραπάνω ζητημάτων λειτουργεί συμπληρωματικά ως προς τη δική μας μέθοδο μιας και υπάρχει αλληλεπίδραση



με τη διαδικασία επικύρωσης της αρχιτεκτονικής. Ο σχεδιασμός της αρχιτεκτονικής πρέπει να αλλάξει σε περίπτωση που δεν μπορούν να ικανοποιηθούν οι απαιτήσεις ποιότητας. Η προτεινόμενη μέθοδος καταδεικνύει τα σχεδιαστικά προβλήματα που αξίζουν προσοχής και πρέπει να προσαρμοστούν ώστε να ικανοποιηθούν οι ποιοτικές απαιτήσεις.

Τα στάδια της προτεινόμενης μεθόδου στα πλαίσια της γενικής διαδικασίας αποτίμησης αρχιτεκτονικών είναι τα παρακάτω:

1. Κατασκευή ενός προσομοιωτή ή πρωτοτύπου που υλοποιεί υποσύνολο της αρχιτεκτονικής.
2. Επικύρωση του πρωτοτύπου ή του προσομοιωτή. Περιλαμβάνει όλα τα απαραίτητα βήματα για να διασφαλιστεί η πιστή αναπαράσταση του προβλήματος. Οι απαραίτητες ενέργειες ποικίλουν ανά πρωτότυπη υλοποίηση ή προσομοιωτή. Στην εργασία [Martis, 2006] παρουσιάζονται ορισμένες από τις ενέργειες για τη διασφάλιση της εγκυρότητας μοντέλων προσομοίωσης.
3. Εκτέλεση πειραμάτων και συλλογή δεδομένων για τη μετέπειτα στατιστική ανάλυση.
4. Εφαρμογή της ανάλυσης παραγόντων για τις ποιοτικές μετρικές στα δεδομένα που συλλέχθηκαν στο προηγούμενο βήμα. Από την ανάλυση προκύπτουν καινούριες μεταβλητές που αντιστοιχούν σε λανθάνοντα χαρακτηριστικά ποιότητας.
5. Ακολουθεί η διαδικασία απόδοσης υπόστασης που αφορά την ερμηνεία των αποτελεσμάτων σε πολυδιάστατο χώρο [Krzpanowski, 1993]. Μπορεί να εφαρμοστεί στην ανάλυση κύριων συνιστωσών και στην ανάλυση παραγόντων όπου εξάγονται νέες λανθάνουσες μεταβλητές. Ερμηνεύονται οι μεταβλητές που προέκυψαν από το προηγούμενο βήμα της μεθόδου.
6. Ακολουθεί η ανάλυση διακύμανσης (ANOVA) των μεταβλητών του τέταρτου βήματος για τη μελέτη και την ερμηνεία των σημείων αντιστάθμισης της αρχιτεκτονικής σε σχέση με τις ποιοτικές μετρικές.

7. Ερμηνεία των αποτελεσμάτων. Με αφετηρία τα αποτελέσματα από την επικύρωση της αρχιτεκτονικής, γίνονται σχεδιαστικές επιλογές για τα αρχιτεκτονικά χαρακτηριστικά που επηρεάζουν τις ποιοτικές μετρικές.

Οι ρόλοι που συμμετέχουν και συνεργάζονται στα διάφορα στάδια της μεθόδου είναι ο αρχιτέκτονας λογισμικού, ο προγραμματιστής της πρωτότυπης υλοποίησης ή ο εμπειρογνώμονας μοντελοποίησης για τη κατασκευή του προσομοιωτή και ο ειδικός στατιστικής ανάλυσης. Ενδέχεται να συμμετέχουν στη διαδικασία επικουρικά και άλλοι, ειδικά κατά την ερμηνεία των αποτελεσμάτων της μεθόδου όταν ο αρχιτέκτονας αποφασίζει για τις απαραίτητες ενέργειες ώστε να επιτευχθούν οι ποιοτικοί στόχοι.

## 5.2 Ανάλυση παραγόντων και λανθάνοντα χαρακτηριστικά ποιότητας

Στην εξεταζόμενη αρχιτεκτονική λογισμικού (βλέπε Ενότητα 4.2), το σημαντικότερο σημείο αντιστάθμισης που επηρεάζει τις μετρικές ποιότητας αφορά το κόστος της ανάκαμψης από αποτυχίες και την απόδοση του συστήματος. Το κόστος της ανάκαμψης δεν είναι άμεσα μετρήσιμο (λανθάνον ποιοτικό χαρακτηριστικό), ενώ η απόδοση εκφράζεται από πολλές μετρικές όπου η κάθε μία εκφράζει διαφορετική πτυχή της αναμενόμενης συμπεριφοράς που εκδηλώνεται κατά την εκτέλεση της υλοποίησης (βλέπε Πίνακα 5.1). Εκτός από τις απεικονιζόμενες μετρικές, καταγράφουμε επιπλέον και τη διαθεσιμότητα των δύο εξυπηρετητών για το εξεταζόμενο φόρτο εργασίας. Οι συναλλαγές ομαδοποιούνται σε κατανεμημένες, τοπικές (όσες εκτελούνται σε ένα μόνο εξυπηρετητή) και τρεις ομάδες συναλλαγών που μοιράζονται τουλάχιστον ένα αντικείμενο. Η συναλλαγές της κλάσης  $tr2$  δε μοιράζονται αντικείμενα με άλλες με αποτέλεσμα να μην συμμετέχει σε κάποια από τις τρεις τελευταίες ομάδες του Πίνακα 5.1.

Η προτεινόμενη μέθοδος αποτελείται από δύο στάδια ανάλυσης που αποσκοπούν στην εύρεση όλων των σημείων αντιστάθμισης που αφορούν τις ποιοτικές μετρικές που μας

	Τύπος Συναλλαγών		tr1, tr3	Κλάσεις Συναλλαγών	
	Κατανεμημένες	Τοπικές		tr4, tr5, tr6	tr7, tr8, tr9
Διαμεταγωγή (ποσοστό επιτυχημένων συναλλαγών)	tput_distr	tput_local	tput_conf0	tput_conf1	tput_conf2
Μέσος χρόνος απόκρισης	response_distr	response_local	response_grp0	response_grp1	response_grp2
Μέσος χρόνος αποκλεισμού	blocking_distr		blocking_grp0	blocking_grp1	blocking_grp2

Πίνακας 5.1: Μετρικές απόδοσης για τον εξεταζόμενο συνθετικό φόρτο εργασίας ενδιαφέρουν. Τα σημεία αντιστάθμισης επιλέγονται από τα αρχιτεκτονικά χαρακτηριστικά με βάση ενδείξεις που παρέχει η ανάλυση παραγόντων σχετικά με τις στατιστικά σημαντικές εξαρτήσεις μεταξύ των ποιοτικών μετρικών.

### 5.2.1 Πειραματικός σχεδιασμός

Σύμφωνα με το τρίτο βήμα της μεθόδου (βλέπε την Ενότητα 5.1), εκτελέσαμε σειρά πειραμάτων με στόχο τη συλλογή δεδομένων που αφορούν τις ποιοτικές μετρικές που παρατίθενται στον Πίνακα 2.1 και την παρατηρούμενη διαθεσιμότητα των δύο εξυπηρετητών. Τα πειράματα για το εξεταζόμενο συνθετικό φόρτο εργασίας εκτελέστηκαν σύμφωνα με τις οδηγίες της εργασίας [Kitchenham et al., 2002]. Τα ερωτήματα που διερευνούμε είναι:

Ερώτημα 1 Ποια είναι τα λανθάνοντα ποιοτικά χαρακτηριστικά που διαμορφώνουν τις εξαρτήσεις μεταξύ των ποιοτικών μετρικών που μας ενδιαφέρουν;

Ερώτημα 2 Ποια αρχιτεκτονικά χαρακτηριστικά είναι σημεία ευαισθησίας και ποια σημεία αντιστάθμισης;

Ερώτημα 3 Δοθέντων των σημείων αντιστάθμισης, ποια αρχιτεκτονικά χαρακτηριστικά έχουν στατιστικά σημαντική επιρροή στα λανθάνοντα ποιοτικά χαρακτηριστικά;

Ερώτημα 4 Πως επηρεάζονται οι ποιοτικές μετρικές από τα αρχιτεκτονικά χαρακτηριστικά με στατιστικά σημαντική επιρροή στα λανθάνοντα ποιοτικά χαρακτηριστικά;

Οι πληροφορίες που αφορούν τον προσομοιωτή και τα εργαλεία στατιστικής ανάλυσης είναι:

1. Η διάρκεια των εκτελέσεων των πειραμάτων προσδιορίστηκε μετά από σειρά δοκιμών ώστε να διασφαλιστεί ότι τα αποτελέσματα συλλέγονται μετά την επίτευξη σταθερής κατάστασης του προσομοιωτή.
2. Σε όλες τις περιπτώσεις φόρτου εργασίας που εξετάστηκαν, διαπιστώθηκε ότι όταν ο προσομοιωμένος χρόνος είναι τουλάχιστον 55 ώρες και 30 λεπτά, αποφεύγεται η εισαγωγή μεροληψίας (bias) στην εκτέλεση του μοντέλου με την επιλογή της κατάστασης έναρξης της προσομοίωσης. Με βάση τα αποτελέσματα της προκαταρκτικής ανάλυσης [Kelton and Law, 1984] ο χρόνος προσομοίωσης προσδιορίστηκε στην παραπάνω τιμή.
3. Ο χρόνος εκτέλεσης των πειραμάτων σε προσωπικό υπολογιστή εφοδιασμένο ένα επεξεργαστή και 1 GB κεντρική μνήμη, κυμάνθηκε από λίγα έως 12 λεπτά ανάλογα με τις παραμέτρους του εκάστοτε πειράματος. Ο ρυθμός άφιξης των συναλλαγών είναι ο σημαντικότερος παράγοντας στο χρόνο εκτέλεσης του πειράματος. Ο προσομοιωτής ACID Sim Tools έχει μικρή κατανάλωση μνήμης (μεταξύ 5 έως 50% της διαθέσιμης μνήμης) και δε παρουσιάζει αστάθειες στη λειτουργία του.
4. Επιλέξαμε το πρόγραμμα SPSS για τη στατιστική ανάλυση των πειραματικών δεδομένων. Τα δεδομένα αφορούν 14 μετρικές απόδοσης (βλέπε Πίνακα 5.1) και δύο μετρικές που δείχνουν τη διαθεσιμότητα των εξυπηρετητών. Συνολικά εξετάστηκαν 16 εξαρτημένες μεταβλητές.

Ο πειραματικός σχεδιασμός βασίστηκε σε προγενέστερες προσομοιώσεις που αποσκοπούσαν στην εξερεύνηση της μεταβλητότητας των ποιοτικών μετρικών στο χώρο των διαθέσιμων σχεδιαστικών επιλογών. Σε όλα τα πειράματα επιλέξαμε το πρωτόκολλο κλειδώματος δύο φάσεων (2PL) για τον έλεγχο της παράλληλης πρόσβασης σε αντικείμενα

διαμοιρασμένα σε δύο ή παραπάνω συναλλαγές. Οι πειραματικοί παράγοντες αντιστοιχούν σε χαρακτηριστικά της εξεταζόμενης αρχιτεκτονικής. Στον Πίνακα 4.4 παρουσιάζουμε τα επιλεγμένα επίπεδα τιμών των πειραματικών παραγόντων. Επιλέχθηκαν οι τιμές όπου οι εξεταζόμενες μετρικές εμφανίζουν σημαντική διακύμανση των τιμών τους. Τα τρία επίπεδα συχνότητάς αποτυχίας των εξυπηρετητών εξετάζουν τη συμπεριφορά του συστήματος όταν οι αποτυχίες προκύπτουν σχετικά συχνά αλλά και σε σχετικά σταθερά συστήματα. Τα δύο επίπεδα συχνότητας άφιξης συναλλαγών για εξυπηρέτηση οδηγούν σε ριζικά διαφορετικό φόρτο εργασίας. Οι αποτυχίες των εξυπηρετητών και οι αφίξεις των συναλλαγών ορίζονται από κατανομές Poisson.

Εκτελέσαμε ένα πείραμα για κάθε συνδυασμό τιμών των πειραματικών παραγόντων του Πίνακα 4.4. Προέκυψαν  $3 * 3 * 3 * 3 * 2 * 3 = 486$  πειράματα. Μία εναλλακτική στρατηγική με σημαντικά λιγότερες απαιτήσεις σε υπολογιστικούς πόρους και λιγότερες εκτελέσεις πειραμάτων είναι ο *ομοιόμορφος πειραματικός σχεδιασμός* (βλέπε [Katsaros et al., 2007] για παράδειγμα χρήσης της μεθόδου).

## 5.2.2 Ανάλυση παραγόντων

Η ανάλυση παραγόντων μειώνει τις διαστάσεις του προβλήματος μέσω του υπολογισμού νέων μεταβλητών, λιγότερες στο πλήθος, οι οποίες διατηρούν το μεγαλύτερο μέρος των πληροφοριών που ενέχονται στις μετρικές των πειραματικών δεδομένων.

Η κεντρική ιδέα είναι να ομαδοποιήσουμε τις 16 εξαρτημένες μεταβλητές, των οποίων οι τιμές προκύπτουν από τη διεξαγωγή πειραμάτων, ώστε οι μεταβλητές με υψηλή συσχέτιση να ενταχθούν στην ίδια ομάδα. Οι ομαδοποιήσεις αντιστοιχούνται σε νέες μεταβλητές επιτυγχάνοντας τη μείωση των διαστάσεων των πειραμάτων δεδομένων. Οι μεταβλητές που προσδιορίζονται από την παραπάνω διαδικασία είναι ασυσχέτιστες και μπορούν να μελετηθούν ανεξάρτητα από τους πειραματικούς παράγοντες.

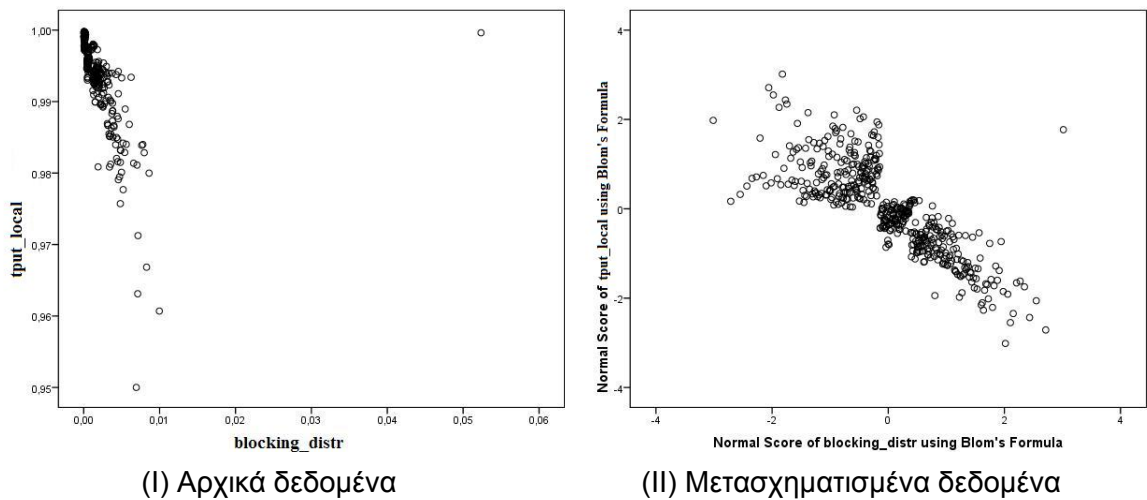
Η ανάλυση παραγόντων αποσκοπεί στον εντοπισμό λανθανουσών μεταβλητών στο αρχικό σύνολο δεδομένων, τις οποίες αποκαλούμε παράγοντες, με βάση τις συσχετίσεις με-

ταξύ των εξαρτημένων μεταβλητών. Για την αποφυγή σύγχυσης με τον όρο “παράγοντες”, όπως χαρακτηρίζονται οι πειραματικοί παράγοντες του Πίνακα 4.4, εφεξής οι λανθάνουσες μεταβλητές θα αποκαλούνται “συστατικά”. Ο χαρακτηρισμός συμβαδίζει με την ωρολόγια που επικρατεί στην ανάλυση παραγόντων.

Βασική προϋπόθεση για την εφαρμογή της ανάλυσης παραγόντων είναι ότι οι τιμές των εξαρτημένων μεταβλητών περιγράφονται από την κανονική κατανομή. Για να την εφαρμογή της μεθόδου απαιτείται μία προεργασία των δεδομένων σύμφωνα με το μετασχηματισμό του Blom [Blom, 1958] ώστε τα δεδομένα να περιγράφονται από την κανονική κατανομή. Ο μετασχηματισμός αξιοποιεί τις διαβαθμίσεις  $r_i$  των  $n$  τιμών και τη συσσωρευτική συνάρτηση της κανονικής κατανομής  $\Phi^{-1}(x)$  (βλέπε Εξίσωση 4.1).

Οι μετασχηματισμένες τιμές των μετρικών περιγράφονται πολύ καλά από την κανονική κατανομή με μέσο 0 και τυπική απόκλιση 1.

Οι συσχετίσεις των αρχικών δεδομένων διατηρούνται και στις μετασχηματισμένες τιμές. Στο Σχήμα 5.2 βλέπουμε τη συσχέτιση των μεταβλητών `trput_local` και `blocking_distr` στα αρχικά δεδομένα (5.2I) και τη συσχέτιση μετά το μετασχηματισμό Blom (5.2II). Επιπλέον, στις μετασχηματισμένες τιμές η συσχέτιση είναι εμφανέστερη.



Σχήμα 5.2: Ισχυρά αρνητική συσχέτιση των μεταβλητών `trput_local` και `blocking_distr` στα αρχικά και τα μετασχηματισμένα δεδομένα

Υποβάλαμε τις μετασχηματισμένες και κανονικοποιημένες εξαρτημένες μεταβλητές στην ανάλυση παραγόντων βασιζόμενοι στη *ανάλυση πρωτευόντων συνιστωσών τύπου Varimax rotation*. Η Varimax rotation αναζητά το γραμμικό συνδυασμό των αρχικών συνιστωσών που μεγιστοποιεί τη διακύμανση. Κάθε εξαρτημένη μεταβλητή συσχετίζεται με μία συνιστώσα ή με ολιγομελή ομάδα συνιστωσών, το οποίο υποβοηθά στην ομαλοποίηση τους. Τελικά, κάθε συνιστώσα αντιστοιχείται σε λίγες μόνο εξαρτημένες μεταβλητές.

Οι καινούριες μεταβλητές υπολογίζονται από τη μέθοδο Anderson–Rubin , που εκτιμά τους συντελεστές των βαθμίδων [Anderson and Rubin, 1956]. Οι υπολογιζόμενες βαθμίδες έχουν μέσο 0 και τυπική απόκλιση 1 και είναι ασυσχέτιστες.

Από την ανάλυση παραγόντων προέκυψαν τρεις συνιστώσες που εξηγούν το 88.81% της παρατηρούμενης διακύμανσης στις 16 αρχικές εξαρτημένες μεταβλητές (μετρικές). Το αποτέλεσμα της μεθόδου είναι η μείωση των διαστάσεων του προβλήματος σε τρεις, αξιοποιώντας τις εσωτερικές συσχετίσεις ανάμεσα στις 16 αρχικές μεταβλητές.

Στο Πίνακα 5.2 δείχνουμε τις τιμές των μεταβλητών αναφορικά με τις τρεις συνιστώσες και τις ομαδοποιήσεις των ποιοτικών μετρικών. Το πρόσημο της τιμής υποδεικνύει τη φύση της συσχέτισης. Δείχνει δηλαδή αν η μεταβλητή συνεισφέρει στην αύξηση της συνιστώσας ή την αντικρούει. Η μεταβλητή `truf_local` για παράδειγμα, ανήκει στην ίδια συνιστώσα με την `blocking_distr` αλλά έχει διαφορετικό πρόσημο λόγω της αρνητικής συσχέτισης μεταξύ των μεταβλητών. Η παραπάνω παρατήρηση μας οδηγεί στο συμπέρασμα ότι *μεγάλα χρονικά διαστήματα αποκλεισμού σε κατανεμημένες συναλλαγές προκαλούν μείωση της διαμεταγωγής των τοπικών συναλλαγών*. Η αιτία της αρνητικής συσχέτισης είναι ότι οι αποκλεισμένες κατανεμημένες συναλλαγές διατηρούν το δικαίωμα αποκλειστικής χρήσης διαμοιρασμένων αντικειμένων με αποτέλεσμα οι τοπικές συναλλαγές να χρειάζονται περισσότερο χρόνο να ολοκληρωθούν. Οι κατανεμημένες συναλλαγές επηρεάζουν τις τοπικές συναλλαγές όλων των εμπλεκόμενων εξυπηρετητών και αυξάνουν σημαντικά την πιθανότητα ακύρωσης των τοπικών συναλλαγών λόγω της λήξης της προθεσμίας ολοκλήρωσής τους.

Οι ομαδοποιήσεις που προέκυψαν από την ανάλυση παραγόντων εμφανίζονται στον Πίνακα 5.3. Η διαδικασία απόδοσης υπόστασης (*reification*) αφορά την ερμηνεία των συνι-

Κανονικοποιημένη Bloom βαθμίδα	Συνιστώσα		
	C1	C2	C3
availability0	<b>-0.800</b>	0.009	0.115
availability1	<b>-0.794</b>	-0.002	0.106
tput_local	<b>-0.831</b>	0.132	0.406
tput_distr	-0.131	-0.125	<b>0.978</b>
tput_confl0	-0.166	0.098	<b>0.955</b>
tput_confl1	-0.138	-0.227	<b>0.950</b>
tput_confl2	-0.140	-0.239	<b>0.947</b>
response_local	0.031	<b>0.724</b>	-0.389
response_distr	0.095	<b>0.953</b>	-0.203
response_grp0	-0.036	<b>0.948</b>	-0.205
response_grp1	-0.017	<b>0.975</b>	0.122
response_grp2	-0.013	<b>0.989</b>	0.040
blocking_distr	<b>0.967</b>	0.062	-0.054
blocking_grp0	<b>0.935</b>	0.068	-0.126
blocking_grp1	<b>0.944</b>	0.007	-0.034
blocking_grp2	<b>0.948</b>	0.038	-0.050

Πίνακας 5.2: Ανεστραμμένος πίνακας συνιστωσών της μεθόδου ανάλυσης παραγόντων στωσών που προέκυψαν από την ανάλυση παραγόντων και απαντά στο πρώτο ερώτημα της Ενότητας 5.2.1. Στο εξεταζόμενο φόρτο εργασίας επεξεργασίας συναλλαγών, τα συμπεράσματα από τις συνιστώσες που προέκυψαν είναι:

- Η συνιστώσα C1 εκφράζει τη διακύμανση των μετρικών που επηρεάζονται κυρίως από το κόστος εγγραφής και ανάγνωσης δεδομένων από μόνιμα μέσα αποθήκευσης

Συνιστώσα	Ποσοστό διακύμανσης	Ομάδα μεταβλητών
C1	35.35%	availability0(-), availability1(-), tput_local(-), blocking_distr(+), blocking_grp0(+), blocking_grp1(+), blocking_grp2(+)
C2	27.64%	response_local(+), response_distr(+), response_grp0(+), response_grp1(+), response_grp2(+)
C3	25.83%	tput_distributed(+), tput_confl0(+), tput_confl1(+), tput_confl2(+)

Πίνακας 5.3: Ομάδες ποιοτικών μετρικών της ανάλυσης παραγόντων



κατά τη διαδικασία δημιουργίας σημείου ελέγχου και ανάνηψης από αποτυχία πτώσης του εξυπηρετητή. Παρατηρούμε ότι οι υψηλές τιμές της συνιστώσας *C1* σχετίζονται με μεγάλα διαστήματα αποκλεισμού των κατανεμημένων συναλλαγών, χαμηλή διαθεσιμότητα των εξυπηρετητών και μικρή διαμεταγωγή των τοπικών συναλλαγών.

- Η συνιστώσα *C2* συνοψίζει τη διακύμανση των μετρικών που αφορούν τους χρόνους απόκρισης των συναλλαγών. Η συσχέτιση των μετρικών με τη συνιστώσα είναι θετική. Αύξηση της τιμής της συνιστώσας σηματοδοτεί αύξηση του χρόνου απόκρισης όλων των συναλλαγών.
- Τέλος, η συνιστώσα *C3* αφορά τις μετρικές διαμεταγωγής και επηρεάζεται κυρίως από καθυστερήσεις στην επικοινωνία των εξυπηρετητών. Επισημαίνουμε ότι σε αντιδιαστολή με της μετρικές διαμεταγωγής κατανεμημένων συναλλαγών, η διαμεταγωγή των τοπικών συναλλαγών εντάσσεται στη συνιστώσα *C1*.

### 5.3 Συνθετικές μετρικές απόδοσης επεξεργασίας συναλλαγών

Είναι δύσκολο να ερμηνευθούν οι τιμές των συνιστωσών αναφορικά με την επίπτωση των ομάδων από μετρικές που εκπροσωπούν στη συνολική συμπεριφορά του συστήματος. Η βασική αιτία της δυσκολίας στην ερμηνεία είναι ότι σε μία συνιστώσα μπορεί να συμμετέχουν παράγοντες που δεν είναι εμπειρικά εμφανής η σχέση που τα συνδέει ούτε και οι μεταξύ τους αλληλεπιδράσεις.

Για τη διευκόλυνση στην ερμηνεία των συνιστωσών είναι επιθυμητό οι κατασκευή τεχνικών μετρικών που θα εξηγούν την παρατηρούμενη διακύμανση, θα επιτρέπουν τη διαισθητική ερμηνεία των υπολογισμένων τιμών και θα υποστηρίζουν την εξαγωγή συμπερασμάτων για τις αρχιτεκτονικές επιλογές.

Οι συνθετικές μετρικές είναι κατάλληλα κατασκευασμένες όταν συνυπολογίζουν τη σχε-

τική σπουδαιότητα του κάθε παράγοντα της συνιστώσας αναφορικά με τους σχεδιαστικούς στόχους ποιότητας. Στη διατύπωση των συνθετικών μετρικών αξιοποιείται το είδος της συσχέτισης μεταξύ των παραγόντων της συνιστώσας (αρνητική ή θετική συσχέτιση) καθώς και ειδικά βάρη στις τιμές ορισμένων παραγόντων για την ενίσχυση ή αποδυνάμωση της επίπτωσης του στη συνθετική μετρική.

Η διαθεσιμότητα των διακομιστών και η διαμεταγωγή των συναλλαγών που υποβάλλονται για εξυπηρέτηση στον κάθε διακομιστή αποτελούν κεντρικούς στόχους βελτίωσης σε αρχιτεκτονικές κατανεμημένων συναλλαγών. Αναφορικά με τους παραπάνω ποιοτικούς στόχους και τη συνιστώσα C1, επιδιώκουμε τον καθορισμό συνθετικής μετρικής που θα αυξάνεται για αρχιτεκτονικές επιλογές που προσεγγίζουν τους στόχους και θα μειώνεται όταν ο στόχος δεν είναι εφικτός. Από την άλλη, επειδή ο χρόνος αποκλεισμού των κατανεμημένων συναλλαγών έχει αρνητική συσχέτιση με τη διαμεταγωγή και τη διαθεσιμότητα, οι επιλογές που αυξάνουν το χρόνο αποκλεισμού πρέπει να οδηγούν σε μείωση της τιμής της συνθετικής μετρικής. Οι παραπάνω απαιτήσεις συνοψίζονται στην Εξίσωση 5.1:

$$component1 = \frac{tput\_local}{blocking\_distr} * \frac{availability0 + availability1}{2} \quad (5.1)$$

Η τιμή του παράγοντα *blocking\_distr* είναι πάντα μεγαλύτερη του μηδενός διότι σε όλα τα εξεταζόμενα πρωτοκολλά οι αποκλεισμοί των κατανεμημένων συναλλαγών είναι αναπόφευκτοι, οπότε η μετρική ορίζεται σε όλες τις περιπτώσεις. Υψηλές τιμές της συνθετικής μετρικής υποδηλώνουν βελτίωση της απόδοσης λόγω της σχέσης μεταξύ των παραγόντων που συμπεριλαμβάνονται στην εξίσωση.

Αν και συνήθως ενδιαφερόμαστε για συγκεκριμένες πτυχές της συμπεριφοράς κατά τη λειτουργία του προγράμματος αναφορικά με κάποια από τα χαρακτηριστικά του φόρτου εργασίας (όπως η χρήση κατανεμημένων πόρων, για παράδειγμα), οι συνθετικές μετρικές συνοψίζουν πληροφορίες που αφορούν συνολικά το φόρτο εργασίας και τα χαρακτηριστικά του. Οπότε, πρέπει να βεβαιωθούμε ότι οι συνθετικές μετρικές δε παρουσιάζουν πληροφορίες διαθέσιμες και από άλλες μετρικές. Στη μετρική που αφορά τη συνιστώσα C2,

λαμβάνουμε υπόψιν τους παράγοντες  $response\_local$  και  $response\_distr$ , αλλά σκοπίμως αγνοούμε τους παράγοντες  $response\_grp0$ ,  $response\_grp1$  και  $response\_grp2$  που συμψηφίζονται στους δύο παράγοντες που αξιοποιούμε. Η εξίσωση της δεύτερης συνθετικής μετρικής είναι:

$$component2 = \frac{response\_local + response\_distr}{2} \quad (5.2)$$

Αντίθετα με την προηγούμενη μετρική, από την Εξίσωση 5.2 συμπεραίνουμε ότι η βελτίωση της αρχιτεκτονικής επιτυγχάνεται για χαμηλές τιμές της συνθετικής μετρικής. Μικρές τιμές αντιστοιχούν σε σύντομη ολοκλήρωση των συναλλαγών. Οι τιμές της μετρικής προέρχονται από το σύνολο  $[0, \infty]$ , όπου ο άπειρος χρόνος σηματοδοτεί την αδυναμία των διακομιστών να αποφανθούν για το αποτέλεσμα της συναλλαγής (ακύρωση ή οριστικοποίηση των αποτελεσμάτων). Στη πραγματικότητα, οι εξυπηρετητές μπορούν πάντα να αποφανθούν για το τελικό αποτέλεσμα οπότε η τιμή της μετρικής δεν μπορεί να προσεγγίσει το άπειρο.

Η συνθετική μετρική για την τρίτη συνιστώσα δίνεται από την εξίσωση:

$$component3 = \frac{tput\_conf10 + tput\_conf11 + tput\_conf12}{3} \quad (5.3)$$

και λαμβάνει τιμές στο διάστημα  $[0, 1]$ , όπου με μηδέν υποδηλώνεται ότι όλες οι συναλλαγές ακυρώνονται ενώ η τιμή ένα δείχνει ότι όλες οι συναλλαγές ολοκληρώνονται με επιτυχία.

Στο Σχήμα 5.3, βλέπουμε τα τρισδιάστατα διαγράμματα διασποράς των τριών συνθετικών μετρικών αναφορικά με κάποιους από τους παράγοντες που καταγράφονται στο Πίνακα 5.2. Σε καθένα από τα διαγράμματα διασποράς, τα σημεία δείχνουν τις τιμές των τριών μετρικών σε σχέση με τις διαβαθμίσεις των τιμών του εξεταζόμενου αρχιτεκτονικού παράγοντα.

Παρατηρούμε μία ξεκάθαρη ομαδοποίηση των σημείων για τις δύο διαβαθμίσεις του παράγοντα MIT (Σχήμα 5.3V). Οι ομάδες των σημείων είναι σχεδόν παράλληλες και η παραπάνω τάση διατηρείται σε όλα τα διαγράμματα. Για  $MIT = 0.4$  παρατηρούμε τη χαμηλό-

τερη τιμή διαμεταγωγής (τρίτη συνθετική μετρική). Συμπεραίνουμε ότι όταν οι συναλλαγές καταφτάνουν με συχνότητα μεγαλύτερη των 0.6 δευτερολέπτων, η διαμεταγωγή των συναλλαγών μειώνεται. Η αιτία της παρατηρούμενης συμπεριφορά είναι ότι πέραν κάποιας συχνότητας άφιξης συναλλαγών, ο ανταγωνισμός για διαμοιρασμένους πόρους (υπολογιστική ισχύ, πρόσβαση σε μόνιμο μέσο αποθήκευσης και διαμοιρασμένα αντικείμενα) ξεπερνά τις δυνατότητες εξυπηρέτησης του εξυπηρετητή και οδηγεί στην ακύρωση μεγάλης μερίδας των συναλλαγών. Επίσης, βλέπουμε ότι για μέσο χρόνος άφιξης συναλλαγών των 0.6 δευτερολέπτων, λόγω της χαμηλότερης κατανάλωσης κοινόχρηστων πόρων, βελτιώνεται ο χρόνος απόκρισης.

Άλλο εμφανές σμήνος σημείων παρατηρείται αναφορικά με τον παράγοντα TT (Σχήμα 5.3IV). Διαπιστώνουμε σύντομη απόκριση σε υποβαλλόμενες συναλλαγές (δεύτερη συνθετική μετρική), όταν ο παράγοντας TT παίρνει την τιμή 0.9 (πράσινο χρώμα) ενώ μεγαλύτερες τιμές των χρονικών περιθωρίων ολοκλήρωσης συναλλαγών προκαλούν σχετική αύξηση των επιτυχημένων συναλλαγών και της διαμεταγωγής τους, επιβαρύνοντας όμως το χρόνο απόκρισης που προσεγγίζει τα διαθέσιμα χρονικά περιθώρια ολοκλήρωσης των συναλλαγών.

Οι μεγαλύτερες τιμές της πρώτης συνθετικής μετρικής σημειώνονται όταν η περίοδος δημιουργίας σημείου ελέγχου είναι 2100 δευτερόλεπτα (κόκκινο χρώμα). Από την άλλη, ο χρόνος ανάνηψης είναι μεγάλος γι' αυτή την περίοδο οπότε η περίοδος των 1300 δευτερολέπτων (μπλε χρώμα) φαντάζει καλύτερη επιλογή. Στις επόμενες ενότητες θα δούμε πως αντιμετωπίζονται συστηματικά οι επιλογές αντιστάθμισης. Η αρνητική επίδραση των συχνών αποτυχιών των εξυπηρετητών (πράσινο χρώμα) στη διαθεσιμότητα και το χρόνο αποκλεισμού των συναλλαγών είναι εμφανής στο Σχήμα 5.3III.

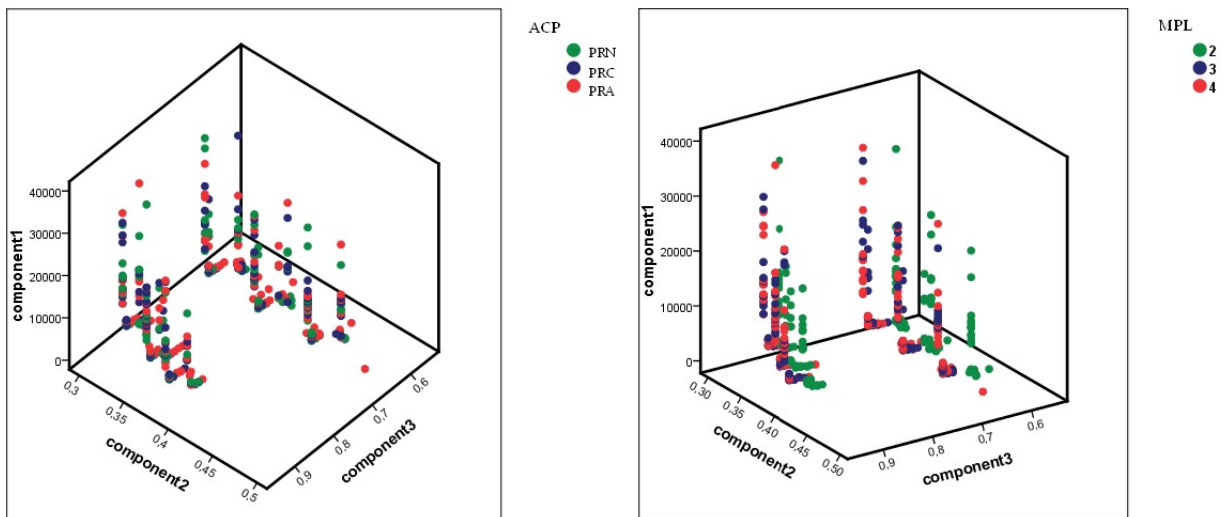
Τέλος, στο Σχήμα 5.3II βλέπουμε ότι για  $MPL = 2$  (πράσινο χρώμα), ο χρόνος απόκρισης των συναλλαγών χειροτερεύει λόγω της καθυστέρησης από την αναμονή για πρόσβαση στους διαθέσιμους υπολογιστικούς πόρους. Τρία ενεργά νήματα επεξεργασίας (μπλε χρώμα) φαίνεται να αξιοποιούν αποτελεσματικότερα τους υπολογιστικούς πόρους χωρίς έντονο ανταγωνισμό για τα αντικείμενα που διαμοιράζονται οι ενεργές συναλλαγές. Περι-

σότερα ενεργά νήματα επεξεργασίας προκαλούν έντονο ανταγωνισμό μεταξύ των ενεργών συναλλαγών και σημειώνονται μεγαλύτεροι χρόνοι απόκρισης.

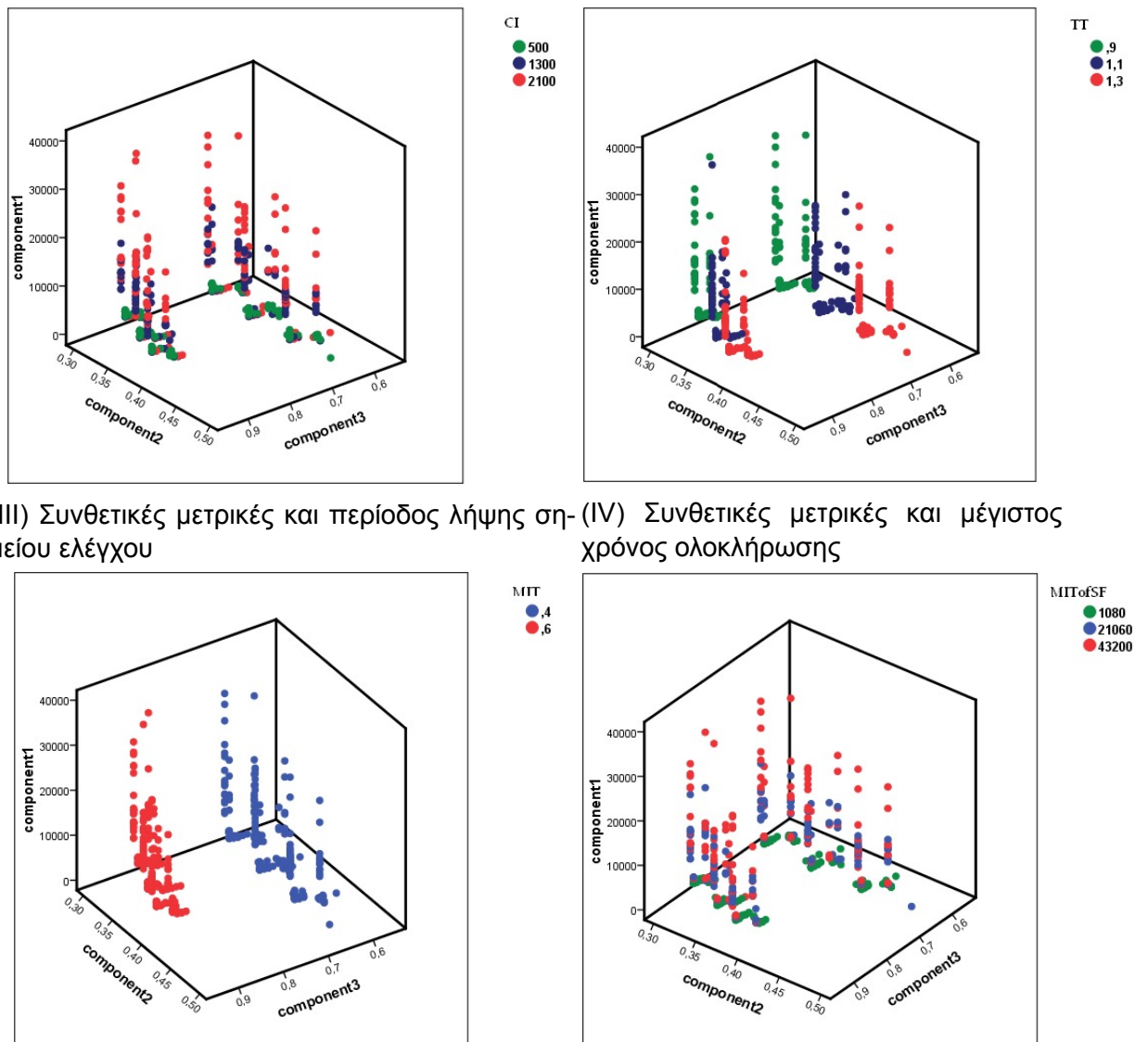
## 5.4 Ανάλυση διακύμανσης για τη διερεύνηση σχέσεων αντιστάθμισης στην απόδοση κατανεμημένων συναλλαγών

Στην Ενότητα 5.2.1 αναφερθήκαμε στα σημεία ευαισθησίας και στις σχέσεις αντιστάθμισης που αποτελούν το ερευνητικό ενδιαφέρον του δεύτερου διατυπωμένου ερωτήματος. Μία πρωταρχική διερεύνηση των σημείων μπορεί να επιτευχθεί από την επισκόπηση διαγραμμάτων θηκοδιαγραμμάτων που απεικονίζουν τη διακύμανση των τριών συνιστωσών αναφορικά με τους παράγοντες. Η Εικόνα 5.4 δείχνει ότι η επίπτωση των παραγόντων είναι διαφορετική για κάθε συνιστώσα.

Η περίοδος δημιουργίας σημείου ελέγχου επιδρά σημαντικά στις ποιοτικές μετρικές της



(I) Συνθετικές μετρικές και πρωτόκολλο συναλλαγών (II) Συνθετικές μετρικές και αριθμός νημάτων επεξεργασίας



Σχήμα 5.3: Επιπτώσεις των πειραματικών παραγόντων στις τρεις συνιστώσες

συνιστώσας C1 (Εικόνα 5.4). Είναι φανερό ότι ο αυτός ο παράγοντες είναι σημείο ευαισθησίας για τη διαθεσιμότητα των εξυπηρετητών, τη διαμεταγωγή των τοπικών συναλλαγών και το χρόνο αποκλεισμού των κατανεμημένων συναλλαγών. Μεταβολές του παράγοντα επιτυγχάνουν ταυτόχρονη βελτίωση (ή χειροτέρευση) των τριών ποιοτικών μετρικών.

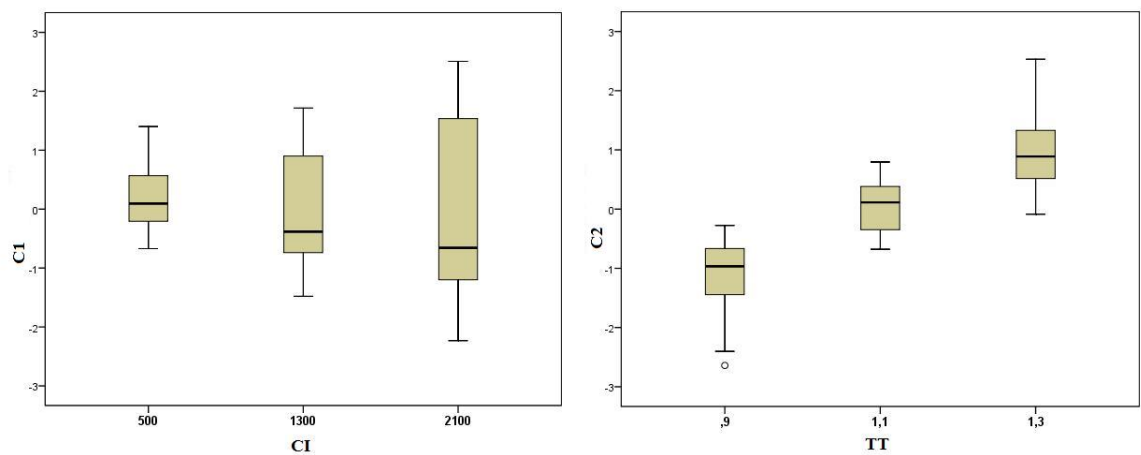
Στις Εικόνες 5.4II και 5.4III παρατηρούμε ότι οι τιμές των συνιστωσών C2 και C3 αυξάνονται για μεγαλύτερα χρονικά περιθώρια ολοκλήρωσης των συναλλαγών. Υψηλότερες τιμές της συνιστώσας C3 σηματοδοτούν την επίτευξη υψηλότερης διαμεταγωγής των συναλλαγών ενώ για τη συνιστώσα C2, οι υψηλές τιμές αντιστοιχούν σε μεγαλύτερους χρόνους απόκρισης (βλέπε τις συσχετίσεις στον Πίνακα 5.2). Το χρονικό όριο ολοκλήρωσης των συναλλαγών επιδρά σε πολλές μετρικές οι οποίες δεν μπορούν να βελτιωθούν ταυτόχρονα με αποτέλεσμα να είναι ένα τυπικό σημείο σχέσης αντιστάθμισης.

### 5.4.1 Ανάλυση διακύμανσης

Χρησιμοποιήσαμε παραγοντική ανάλυση διακύμανσης (ANOVA) στις τρεις συνιστώσες για την καλύτερη κατανόηση των σημείων σχέσεων αντιστάθμισης και για να απαντήσουμε στο τρίτο ερώτημα της έρευνας όπως διατυπώθηκε στην Ενότητα 5.2.1. Η ANOVA αξιοποιείται στη μελέτη της επίδρασης που έχουν οι παράγοντες σε συνεχείς εξαρτημένες μεταβλητές πολυπαραγοντικού πειράματος. Στη δική μας ανάλυση, θεωρούμε την κάθε συνιστώσα ως εξαρτημένη μεταβλητή και διεξάγουμε την ANOVA για όλους τους ανά δύο συνδυασμούς των συνιστωσών με τους έξι πειραματικούς παράγοντες του Πίνακα 4.4.

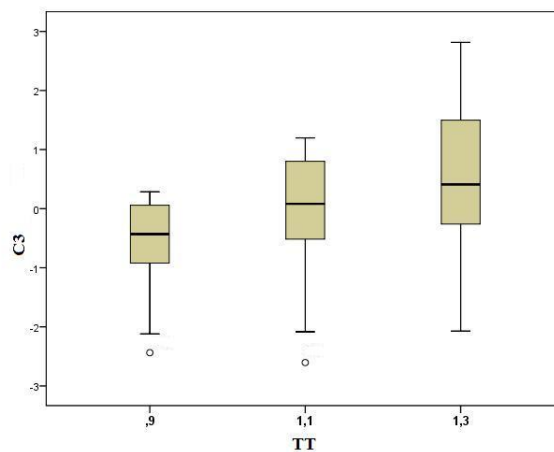
Η ANOVA επιτυγχάνει την κατασκευή αυξητικού μοντέλου που περιέχει τις κύριες επιδράσεις των παραγόντων και τις μεταξύ τους αλληλεπιδράσεις. Υπάρχουν πολλοί διαθέσιμοι στατιστικοί έλεγχοι και μετρικές που βοηθούν στην εξαγωγή συμπερασμάτων σχετικά με τις σχέσεις μεταξύ των παραγόντων και των εξαρτημένων μεταβλητών (βλέπε [Field, 2009] για λεπτομερή ανασκόπηση των διαθέσιμων ελέγχων και την υλοποίησή τους στο στατιστικό εργαλείο SPSS). Στη δική μας ανάλυση αξιοποιήσαμε τους παρακάτω ελέγχους:

- Οι παράγοντες με σημαντικές επιπτώσεις στις εξαρτημένες μεταβλητές αποτιμώνται από στατιστική δοκιμή F. Θεωρούμε στατιστικά σημαντικές τις επιπτώσεις των παραγόντων ή τις αλληλεπιδράσεις τους που αποτιμώνται σε 0,05 από το στατιστικά έλεγχο.



(I) Συνιστώσα C1 και περίοδος δημιουργίας σημείου ελέγχου

(II) Συνιστώσα C2 και χρονικό όριο ολοκλήρωσης συναλλαγών



(III) Συνιστώσα C3 και χρονικό όριο ολοκλήρωσης συναλλαγών

Σχήμα 5.4: Επιπτώσεις των πειραματικών παραγόντων στις τρεις συνιστώσες

- Το ποσοστό της διακύμανσης των εξαρτημένων μεταβλητών που αποδίδεται στους παράγοντες του μοντέλου προσδιορίζεται από τη R τετραγωνισμένη τιμή στο διάστημα  $[0, 1]$ .
- Με την H μερικά τετραγωνισμένη τιμή προσδιορίζουμε τους παράγοντες με τις σημαντικότερες επιδράσεις στις εξαρτημένες μεταβλητές. Για κάθε παράγοντά ή αλληλεπίδραση, η H τετραγωνισμένη τιμή δείχνει το βαθμό συσχέτισης του παράγοντα με



την εξαρτημένη μεταβλητή.

- Οι δοκιμές πολλών συγκρίσεων εντοπίζουν στατιστικά σημαντικές διαφορές μεταξύ των τιμών του εκάστοτε πειραματικού παράγοντα σε σχέση με τις εξαρτημένες μεταβλητές.
- Τα διαγράμματα αλληλεπίδρασης αποκαλύπτουν τις σημαντικότερες αλληλεπιδράσεις των παραγόντων. Είναι γραμμικό διάγραμμα όπου κάθε συντεταγμένη ορίζει την εκτιμώμενη περιθώρια μέση τιμή (marginal mean) της εξαρτημένης μεταβλητής σε κάποια τιμή του παράγοντα. Οι τιμές κάποιου δευτέρου παράγοντα απεικονίζονται για το σχηματισμό γραμμών στη γραφική παράσταση. Τέλος, για κάθε δυνατή τιμή κάποιου τρίτου παράγοντα, προκύπτει άλλη γραφική παράσταση.

Δοκιμάσαμε στην ανάλυσή μας για κάθε συνιστώσα (εξαρτημένη μεταβλητή) διάφορα μοντέλα συμπεριλαμβανομένου του μοντέλου με τους έξι παράγοντες, όλους τους πιθανούς συνδυασμούς μη διατεταγμένων διόδων των παραγόντων (15 ανά δύο αλληλεπιδράσεις) και τους συνδυασμούς τριάδων παραγόντων (20 ανά τρεις αλληλεπιδράσεις παραγόντων). Αλληλεπιδράσεις που αφορούν περισσότερο από τρεις παράγοντες δε λαμβάνονται υπόψιν διότι η ερμηνεία του στατιστικού αποτελέσματος είναι εξαιρετικά δύσκολη. Παρουσιάζουμε τα αποτελέσματα που αφορούν τους παράγοντες και τις αλληλεπιδράσεις με τιμή F-test μικρότερη από 0.05.

### **ANOVA για το συστατικό C1**

Οι στατιστικά σημαντικές επιπτώσεις των παραγόντων και των αλληλεπιδράσεων στη συνιστώσα C1, παρουσιάζονται στον Πίνακα 5.4. Η τιμή της R-squared υποδεικνύει ότι το συγκεκριμένο μοντέλο ANOVA περιγράφει το 93% της διακύμανσης της συνιστώσας C1. Από τη μερική eta-squared οδηγούμαστε στο συμπέρασμα ότι οι σημαντικότερες αλληλεπιδράσεις είναι ο παράγοντας MITofSF (92.1%), η αλληλεπίδραση CI \* MITofSF (71.2%), και οι παράγοντες CI (21.3%), MIT (14.5%) και TT (10.5%). Οι υπόλοιπες αλληλεπιδρά-

Παράγοντας	F-test	Σημαντικότητα	Μερική eta-square
ACP	5.886	0.003	0.025
CI	62.643	0.000	0.213
TT	27.215	0.000	0.105
MIT	78.210	0.000	0.145
MITofSF	2696.656	0.000	0.921
CI * MITofSF	284.876	0.000	0.712
CI * MIT	4.425	0.012	0.019
MIT * MITofSF	4.635	0.010	0.020
ACP * MITofSF	3.499	0.008	0.029
TT * MIT	8.945	0.000	0.037

ACP	Πρωτόκολλο οριστικοποίησης συναλλαγών
MPL	Ενεργά νήματα επεξεργασίας
CI	Περίοδος δημιουργίας σημείου ελέγχου
TT	Χρονικό περιθώριο ολοκλήρωσης συναλλαγών
R-square	0.937
Προσαρμοσμένη R-square	0.934

Πίνακας 5.4: Στατιστικά σημαντικές επιδράσεις της συνιστώσας C1

σεις, παρόλο που έχουν στατιστικά σημαντικές επιπτώσεις, περιγράφουν μικρό μέρος της διακύμανσης (λιγότερο από 5%) της συνιστώσας C1.

Τα αποτελέσματα από δοκιμές post hoc, παρουσιάζονται στον Πίνακα 5.4. Οι τιμές των παραγόντων MIT και MITofSF παραλείπονται μιας και μας ενδιαφέρουν μόνο εκείνοι οι παράγοντες που αποτελούν αρχιτεκτονικές επιλογές. Στον Πίνακα σημειώνονται τα ζεύγη των τιμών παραγόντων που παρουσιάζουν στατιστικά σημαντικές διαφορές. Η περίοδος λήψης σημείων ελέγχου 500 δευτερολέπτων διαφέρει σημαντικά από τις άλλες δύο τιμές (1300 και 2100 δευτερόλεπτα) σχετικά με τις επιπτώσεις στη συνιστώσα C1. Το ίδιο παρατηρείται για όλες τις διαφορετικές τιμές του παράγοντα TT.

Το διάγραμμα της Εικόνας 5.5 δείχνει την αλληλεπίδραση μεταξύ των παραγόντων CI και MITofSF, που αντιστοιχούν στο 71.2% της διακύμανσης της συνιστώσας C1. Το διάγραμμα αφορά το πρωτόκολλο 2PC PRN, και είναι παρόμοιο με τα αντίστοιχα διαγράμματα

		500	1300	2100
CI	500		•	•
	1300	•		
	2100	•		
		0.9	1.1	1.3
TT	0.9		•	•
	1.1	•		•
	1.3	•	•	

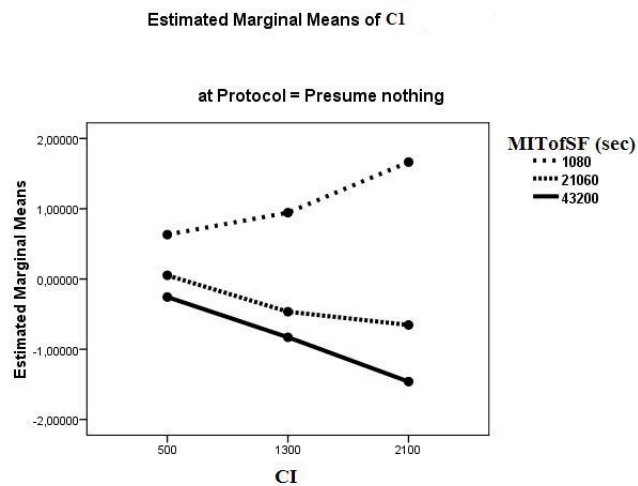
CI Περίοδος δημιουργίας σημείου ελέγχου

TT Χρονικό περιθώριο ολοκλήρωσης συναλλαγών

Πίνακας 5.5: Στατιστικά σημαντικές διαφορές μεταξύ των παραγόντων της συνιστώσας C1

για τα πρωτόκολλα 2PC PRC και 2PC PRA που επιλέξαμε να μην τα παρουσιάσουμε. Η ομοιότητα των διαγραμμάτων οφείλετε στο ότι τα πρωτόκολλα δεν αλληλεπιδρούν με τους δύο παράγοντες.

Οι συχνές λήψεις σημείων ελέγχου μειώνουν σημαντικά το απαιτούμενο χρόνο ανάνηψης των εξυπηρετητών από αποτυχίες αλλά παράλληλα αυξάνεται το κόστος εγγραφής ή



Σχήμα 5.5: Οι αλληλεπιδράσεις μεταξύ των παραγόντων CI και MITofSF στη συνιστώσα C1

ανάγνωσης από μόνιμα μέσα αποθήκευσης λόγω του ανταγωνισμού με τη διαδικασία λήψης σημείου ελέγχου. Όταν οι αποτυχίες πτώσης του συστήματος εμφανίζονται με μικρή συχνότητα, η επιβάρυνση κατά την κανονική λειτουργία του εξυπηρετητή είναι σημαντικά μεγαλύτερη από το όφελος της ταχύτερης ανάνηψης. Από την άλλη, όταν οι αποτυχίες είναι συχνές, η ταχύτερη ανάνηψη βελτιώνει τη διαθεσιμότητα των εξυπηρετητών και το χρόνο αποκλεισμού των κατανεμημένων συναλλαγών, δικαιολογώντας το κόστος κατά την κανονική λειτουργία. Οι παραπάνω αλληλεξαρτήσεις είναι εμφανείς στο Σχήμα 5.5.

### **ANOVA για το συστατικό C2**

Οι επιδράσεις στη συνιστώσα C2, συνοψίζει τις μετρικές που αφορούν τους χρόνους απόκρισης, που παρουσιάζουν στατιστική σημασία, εμφανίζονται στον Πίνακα 5.6. Η τιμή της R-squared υποδεικνύει ότι το 99.5% της διακύμανσης της συνιστώσας C2 περιγράφεται από το μοντέλο. Παρατηρούμε ότι όλοι οι παράγοντες έχουν σημαντικές επιδράσεις. Σημαντικές επιδράσεις εμφανίζουν και οι αλληλεπιδράσεις δύο και τριών παραγόντων. Οι σημαντικότερες επιδράσεις, όπως καταδεικνύονται από τη μερική eta-square, είναι οι παράγοντες TT (99.3%), MIT (96.5%), MPL (95.1%) και οι μεταξύ τους αλληλεπίδραση TT \* MIT \* MPL (60.1%). Η επιλογή του πρωτόκολλου φαίνεται να έχει σημαντική επίπτωση είτε από μόνος του (40.4%) είτε σε συνδυασμό με άλλους παράγοντες.

Στον Πίνακα 5.7 βλέπουμε τα αποτελέσματα των post hoc δοκιμών που μας οδηγούν στο συμπέρασμα ότι το πρωτόκολλο 2PC PRA εμφανίζει σημαντικές διαφορές σε σχέση με το χρόνο απόκρισης σε συγκριτικά με τα άλλα δύο πρωτόκολλα. Τα αποτελέσματα που αφορούν τους παράγοντες MPL και TT, δείχνουν ότι όλες σε όλα τα επίπεδα τιμών υπάρχουν σημαντικές διαφορές στην επίπτωση που αφορά τη συνιστώσα C2, όπως υποδεικνύει και το Σχήμα 5.4II.

Η αλληλεπίδραση μεταξύ των τριών παραγόντων MPL, TT και MIT παρουσιάζεται στο Σχήμα 5.6. Είναι η σημαντικότερη αλληλεπίδραση που αφορά τη συνιστώσα C2. Από τα δύο διαγράμματα παρατηρούμε μεγαλύτερους χρόνους απόκρισης για  $MIT = 0.4$  (5.6I)

Παράγοντας	F-test	Σημαντικότητα	Μερική eta-square
ACP	130.303	0.000	0.404
MPL	3761.003	0.000	0.951
CI	18.469	0.000	0.088
TT	26339.988	0.000	0.993
MIT	10540.846	0.000	0.965
MITofSF	37.216	0.000	0.162
CI * MITofSF	5.003	0.001	0.050
MIT * MITofSF	11.016	0.000	0.054
ACP * MITofSF	41.395	0.000	0.301
TT * MITofSF	3.616	0.007	0.036
TT * MIT	84.863	0.000	0.307
ACP * TT	9.833	0.000	0.093
MPL * TT	88.956	0.000	0.481
CI * MIT * MITofSF	4.262	0.000	0.062
MPL * CI * MITofSF	2.227	0.010	0.065
ACP * MPL * MITofSF	2.438	0.014	0.048
ACP * TT * MITofSF	4.526	0.000	0.086
ACP * MPL * MIT	4.241	0.002	0.042
ACP * TT * MIT	2.611	0.035	0.026
MPL * TT * MIT	144.837	0.000	0.601
ACP * MPL * TT	5.764	0.000	0.107
ACP	Πρωτόκολλο οριστικοποίησης συναλλαγών		
MPL	Ενεργά νήματα επεξεργασίας		
CI	Περίοδος δημιουργίας σημείου ελέγχου		
TT	Χρονικό περιθώριο ολοκλήρωσης συναλλαγών		
R-square	0.995		
Προσαρμοσμένη R-square	0.993		

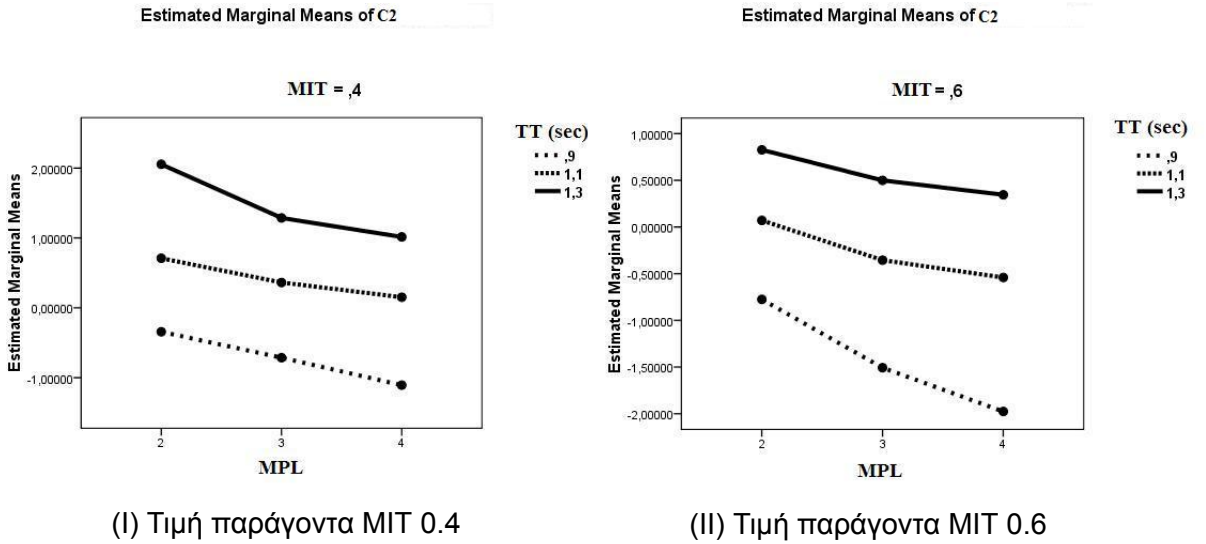
Πίνακας 5.6: Στατιστικά σημαντικές επιδράσεις της συνιστώσας C2 (μετρικές χρόνου απόκρισης)

σε σχέση με την τιμή 0.6 του ίδιου παράγοντα ( 5.6II). Επίσης, παρατηρούμε την τάση για απότομη μείωση του χρόνου απόκρισης όταν το επίπεδο πολυπρογραμματισμού από 2 σε 4 ενεργά νήματα επεξεργασίας, έχοντας  $MIT = 0.6$  και  $TT = 0.9$ . Συμπεραίνουμε ότι η αύξηση του αριθμού των ενεργών νημάτων, με τις προαναφερθέντες τιμές των υπολοίπων

		2PC PRN	2PC PRC	2PC PRA
APC	2PC PRN			•
	2PC PRC			•
	2PC PRA	•	•	
		2	3	4
MPL	2		•	•
	3	•		•
	4	•	•	
		0.9	1.1	1.3
TT	0.9		•	•
	1.1	•		•
	1.3	•	•	

ACP Πρωτόκολλο οριστικοποίησης συναλλαγών  
 MPL Ενεργά νήματα επεξεργασίας  
 TT Χρονικό περιθώριο ολοκλήρωσης συναλλαγών

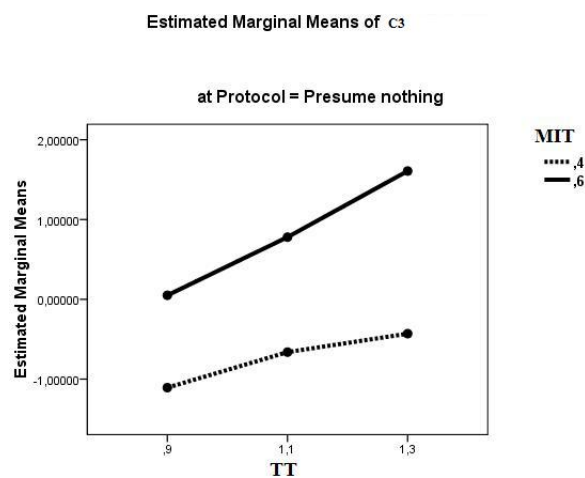
Πίνακας 5.7: Στατιστικά σημαντικές διαφορές μεταξύ των παραγόντων της συνιστώσας C2 (μετρικές χρόνου απόκρισης)



Σχήμα 5.6: Οι αλληλεπιδράσεις μεταξύ των παραγόντων MPL, TT και MIT στη συνιστώσα C2

παραγόντων, είναι αποτελεσματικότερη στρατηγική για την επίτευξη ταχύτερης απόκρισης απ' ότι για άλλες πιθανές τους τιμές. Άλλη μία περίπτωση βελτίωσης του χρόνου απόκρισης με  $MPL = 3$ , σημειώνεται όταν έχουμε συχνές αφίξεις συναλλαγών ( $MIT = 0.4$ ) και το μέγιστο περιθώριο ολοκλήρωσης της επεξεργασίας των συναλλαγών ( $TT = 1.3$ ).

### ANOVA για το συστατικό C3



Σχήμα 5.7: Οι αλληλεπιδράσεις μεταξύ των παραγόντων TT και MIT στη συνιστώσα C3

Βλέπουμε στον Πίνακα 5.8 τα αποτελέσματα που αποκομίσθηκαν από το μοντέλο της ANOVA για τη συνιστώσα C3. Το μοντέλο περιγράφει το 99.1% της παρατηρούμενης διακύμανσης της συνιστώσας. Όλοι οι παράγοντες έχουν σημαντικές επιπτώσεις καθώς και αλληλεπιδράσεις. Με βάση τη μερική  $\eta^2$ -squared, οι σημαντικότερες επιδράσεις οφείλονται στους παράγοντες MIT (98.5%), TT (95.7%), MPL (91.0%), MITofSF (65.8%) και την αλληλεπίδραση των παραγόντων MIT και MITofSF.

Από τις post hoc δοκιμές (βλέπε Πίνακα 5.9) διαπιστώθηκε πως για το πρωτόκολλο 2PC PRA, οι μετρικές είναι σημαντικά διαφοροποιημένες σε σχέση με τις άλλες δύο επιλογές. Στους παράγοντες MPL, CI και TT παρατηρούνται σημαντικές διαφορές σε όλες τις διαθέσιμες τιμές.

Παράγοντας	F-test	Σημαντικότητα	Μερική eta-square
ACP	48.171	0.000	0.207
MPL	1869.720	0.000	0.910
CI	57.664	0.000	0.239
TT	4126.011	0.000	0.957
MIT	24767.104	0.000	0.985
MITofSF	353.506	0.000	0.658
CI * MITofSF	52.601	0.000	0.364
CI * MIT	12.842	0.000	0.065
ACP * CI	5.848	0.000	0.060
MPL * CI	3.535	0.008	0.037
CI * TT	12.760	0.000	0.122
MIT * MITofSF	4.755	0.009	0.025
ACP * MITofSF	46.461	0.000	0.336
MPL * MITofSF	6.558	0.000	0.067
TT * MITofSF	43.777	0.000	0.322
MPL * MIT	311.983	0.000	0.629
TT * MIT	568.012	0.000	0.755
MPL * TT	11.876	0.000	0.114
CI * MIT * MITofSF	2.514	0.041	0.027
MPL * CI * MITofSF	2.959	0.003	0.060
CI * TT * MITofSF	7.165	0.000	0.135
ACP * CI * MIT	2.629	0.034	0.028
MPL * CI * MIT	3.919	0.004	0.041
CI * TT * MIT	3.283	0.012	0.034
ACP * MIT * MITofSF	6.403	0.000	0.065
MPL * MIT * MITofSF	16.472	0.000	0.152
TT * MIT * MITofSF	35.348	0.000	0.278
ACP * MPL * MIT	2.070	0.038	0.043
ACP * TT * MIT	3.220	0.001	0.065
MPL * TT * MIT	14.963	0.000	0.140

ACP	Πρωτόκολλο οριστικοποίησης συναλλαγών
MPL	Ενεργά νήματα επεξεργασίας
CI	Περίοδος δημιουργίας σημείου ελέγχου
TT	Χρονικό περιθώριο ολοκλήρωσης συναλλαγών
R-square	0.991
Προσαρμοσμένη R-square	0.988

Πίνακας 5.8: Στατιστικά σημαντικές επιδράσεις της συνιστώσας C3



		2PC PRN	2PC PRC	2PC PRA
APC	2PC PRN			•
	2PC PRC			•
	2PC PRA	•	•	
		2	3	4
MPL	2		•	•
	3	•		•
	4	•	•	
		500	1300	2100
CI	500		•	•
	1300	•		•
	2100	•	•	
		0.9	1.1	1.3
TT	0.9		•	•
	1.1	•		•
	1.3	•	•	

ACP Πρωτόκολλο οριστικοποίησης συναλλαγών  
 MPL Ενεργά νήματα επεξεργασίας  
 CI Περίοδος δημιουργίας σημείου ελέγχου  
 TT Χρονικό περιθώριο ολοκλήρωσης συναλλαγών

Πίνακας 5.9: Στατιστικά σημαντικές διαφορές μεταξύ των παραγόντων της συνιστώσας C3

Η αλληλεπίδραση μεταξύ των παραγόντων TT και MIT βρέθηκε να έχει τη σημαντικότερη επίπτωση στη συνιστώσα C3. Το διάγραμμα του Σχήματος 5.7 δείχνει τη αλληλεπίδραση αναφορικά με το πρωτόκολλο 2PC PRN. Για τα άλλα δύο πρωτόκολλα έχουμε παρόμοια διαγράμματα διότι το πρωτόκολλο δεν αλληλεπιδρά με τους άλλους δύο παράγοντας. Αξίζει να σημειωθεί ότι για  $MIT = 0.6$  παρατηρείται μεγαλύτερη βελτίωση της συνιστώσας C3, όταν μεταβάλλεται η τιμή του παράγοντα TT από 0.9 σε 1.3, απ' ό,τι όταν η τιμή του MIT είναι 0.4. Από τις παραπάνω παρατηρήσεις οδηγούμαστε στο συμπέρασμα ότι η αύξηση των χρονικών περιθωρίων ολοκλήρωσης των συναλλαγών δεν είναι ιδιαίτερα αποτελεσματική στη βελτίωση της συνιστώσας, όταν έχουμε συχνές αφίξεις νέων συναλλαγών.

Άλλες ενδιαφέρουσες επιδράσεις που βρέθηκαν από άλλα διαγράμματα αλληλεπίδρασης που τα παραλείπουμε χάριν συντομίας, συμπεριλαμβάνουν ένα σημείο ευαισθησίας που αφορά την επιλογή του πρωτοκόλλου ACP και τις μετρικές της συνιστώσας C1 καθώς και σημείο ευαισθησίας της μετρικής C3 αναφορικά με τον παράγοντα C1.

Συμπερασματικά, από τα μοντέλα ANOVA έγινε εφικτός ο προσδιορισμός στατιστικά σημαντικών επιδράσεων των παραγόντων και η ερμηνεία των επιπτώσεων των παραγόντων και των αλληλεπιδράσεων στις ποιοτικές μετρικές. Ερευνήσαμε όλες τις σημαντικές αλληλεπιδράσεις και εντοπίσαμε πιθανούς κινδύνους που παρουσιάζουν κάποιοι από τους συνδυασμούς τιμών των παραγόντων για την επίτευξη των στόχων ποιότητας της αρχιτεκτονικής.

Από την ανάλυση προέκυψαν συμπεράσματα που απαντούν στο τέταρτο ερώτημα της Ενότητας 5.2.1.

#### 5.4.2 Γραφική απεικόνιση σχέσεων αντιστάθμισης

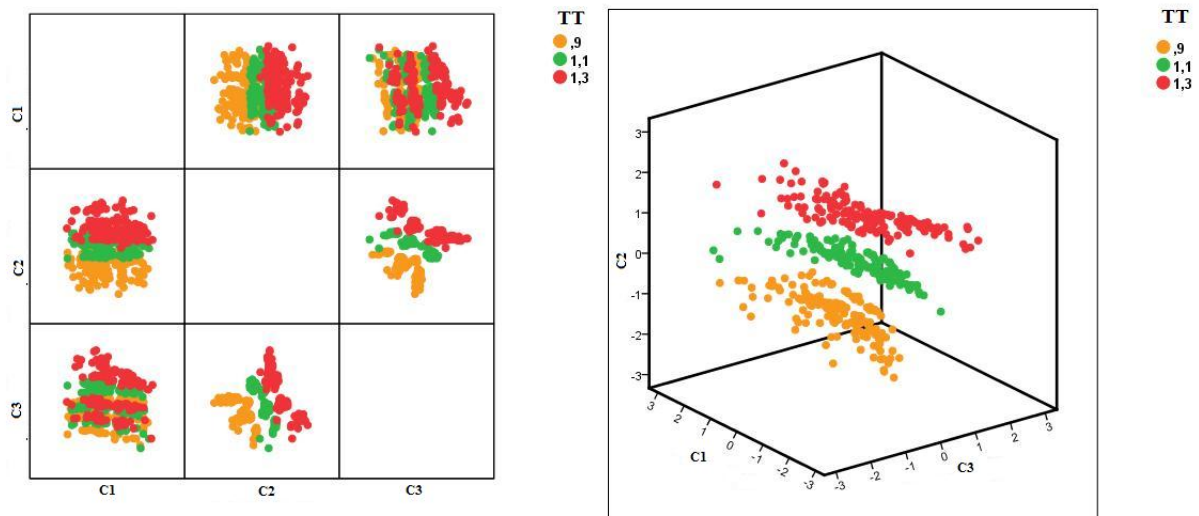
Η μείωση των 16 αρχικών παραγόντων σε 3 συνιστώσες μας επιτρέπει τη προβολή των πειραματικών σημείων σε τρισδιάστατο διάγραμμα διασποράς. Το διάγραμμα βοηθά στην εξαγωγή αρχικών συμπερασμάτων σχετικά με την επίπτωση των αρχιτεκτονικών επιλογών στις ποιοτικές μετρικές. Στα διαγράμματα του Σχήματος 5.8 δείχνουμε και τις τρεις συνιστώσες ενώ τα σημεία που αφορούν διαφορετικές τιμές του απεικονιζόμενου παράγοντα διακρίνονται από διαφορετικό χρώμα το καθένα.

Διαγράμματα όπως εκείνα του Σχήματος 5.8 μας βοηθούν στον εντοπισμό παραγόντων με ισχυρή επίδραση στο σύνολο των μετρήσεων που απεικονίζονται στο διάγραμμα, παρατηρώντας για σχηματισμό ομάδων από σημεία σε κάποιες από τις τιμές των παραγόντων. Συγκεκριμένα, στο Σχήμα 5.8I δείχνουμε το διάγραμμα πίνακα που αφορά τα ζεύγη των τριών συνιστωσών και στο Σχήμα 5.8II βλέπουμε το τρισδιάστατο διάγραμμα διασποράς των τριών συνιστωσών αναφορικά με τον παράγοντα TT.

Από τα διαγράμματα βλέπουμε ξεκάθαρα την ομαδοποίηση των σημείων, όπου τα ση-

μεία που αφορούν την τιμή  $TT = 0.9$  αντιστοιχούν σε σχετικά χαμηλές τιμές χρόνου απόκρισης των συναλλαγών (C2) και χαμηλή τιμή των μετρικών διαμεταγωγής της συνιστώσας C3, οι οποίες επηρεάζονται κυρίως από υστερήσεις στην επικοινωνία. Παρατηρούμε αυτά τα αποτελέσματα διότι σημαντικό μέρος των συναλλαγών δεν ολοκληρώνονται σε χρόνο μικρότερο των 0.9 δευτερολέπτων και ακυρώνονται. Μόνο οι συναλλαγές που ολοκληρώνονται με επιτυχία στα δεδομένα χρονικά όρια συνεισφέρουν στις μετρικές μέσου χρόνου απόκρισης και διαμεταγωγής συναλλαγών.

Παρόμοιες ομαδοποιήσεις των σημείων βρέθηκαν και για τον παράγοντα MIT, όπου τα σημεία για  $MIT = 0.4$  σχηματίζουν ομάδα που αντιστοιχεί σε σχετικά υψηλές τιμές της συνιστώσας C2 (μεγάλοι χρόνοι απόκρισης) και χαμηλές τιμές της συνιστώσας C3. Το παρατηρούμενο αποτέλεσμα δεν εκπλήσσει. Μεγάλο μέρος των συναλλαγών που φτάνουν με ρυθμό  $MIT = 0.4$  τελικά ακυρώνονται οπότε βρίσκουμε χαμηλή διαμεταγωγή για τις κατανομημένες συναλλαγές και για τις ομάδες συναλλαγών που μοιράζονται ένα ή



(I) Διάγραμμα πίνακα των ζευγών συνιστωσών και το παράγοντα TT

(II) Τρισδιάστατο διάγραμμα διασποράς της επίδρασης του παράγοντα TT στις τρεις συνιστώσες

Σχήμα 5.8: Διαγράμματα για την επίδραση του παράγοντα TT στις τρεις συνιστώσες

παραπάνω αντικείμενα όπου μπορεί να έχουμε αδιέξοδα και αμοιβαίες ακυρώσεις. Επιπλέον, από τα διαγράμματα διασποράς βρέθηκε ότι για συχνές αποτυχίες των εξυπηρετητών ( $MITofSF = 1080$ ), τα αντίστοιχα σημεία σχηματίζουν ξεχωριστή ομάδα. Η ακραία τιμή του παράγοντα αντανακλάται στα διαγράμματα που αποκομίσαμε. Στη πραγματικότητα, τόσες συχνές αποτυχίες θα οδηγούσαν στην αντικατάσταση του συστήματος επεξεργασίας συναλλαγών με άλλο πιο αξιόπιστο.

### 5.4.3 Απειλές στην εγκυρότητα των αποτελεσμάτων

Στις προηγούμενες ενότητες δείξαμε την αποτελεσματικότητα της μεθόδου στον εντοπισμό σημείων αντιστάθμισης και στην εκτίμηση των σημαντικότερων επιδράσεων των παραγόντων ή συνδυασμών παραγόντων, στις ποιοτικές μετρικές. Σε αυτή την Ενότητα διερευνούμε συνθήκες οι οποίες μπορεί να μας οδηγήσουν σε λανθασμένα αποτελέσματα.

Οφείλουμε να εξετάσουμε με ιδιαίτερη προσοχή την ύπαρξη σύμμεικτων παραγόντων (confounding factor). Πρόκειται για παράγοντες που συμμεταβάλλονται με παράγοντες της αρχιτεκτονικής που δε περιλαμβάνονται στο πειραματικό σχεδιασμό και στην επακόλουθη ανάλυση αλλά επιδρούν στις εξαρτημένες μεταβλητές (ποιοτικές μετρικές).

Τέτοιοι παράγοντες δε βρέθηκαν στην ανάλυση της αρχιτεκτονικής κατανεμημένων συναλλαγών. Ας υποθέσουμε ωστόσο ότι οι πόροι που διατίθενται για τη διαχείριση της ταυτόχρονης πρόσβασης σε πόρους κοινόχρηστους ανάμεσα στις ενεργές συναλλαγές, είχαν σημαντική επίπτωση στις μετρήσεις των ποιοτικών μετρικών. Σε αυτή την περίπτωση, τα αποτελέσματα της ανάλυσης δε θα ήταν ακριβή, αφού μέρος της παρατηρούμενης διακύμανσης των τιμών θα αποδίδονταν λανθασμένα στους συμμετάβλητους παράγοντες.

Τα δεδομένα της ανάλυσης που παρουσιάσαμε προέρχονται από προσομοίωση. Όταν όμως τα δεδομένα προέρχονται από μετρήσεις (όπως από πρωτότυπη υλοποίηση λόγου χάρη), υπάρχουν ζητήματα που χρίζουν ιδιαίτερης προσοχής για να αποτραπεί κάθε απειλή προς την εσωτερική εγκυρότητα της ανάλυσης. Στα δεδομένα από μετρήσεις, παράγοντες όπως το κόστος που ενέχει η λήψη των μετρήσεων και η παρακολούθηση και καταγραφή

των δραστηριοτήτων της υλοποίησης πρέπει να συμπεριλαμβάνονται στον πειραματικό σχεδιασμό και στην ανάλυση.

Η γενίκευση των συμπερασμάτων της ανάλυσης σχετικά με κάποια αρχιτεκτονική, δεν μπορούν να γενικευτούν σε παραλλαγές της και σε άλλες αρχιτεκτονικές. Η λανθασμένη γενίκευση αποτελεί εξωτερική απειλή της εγκυρότητα των συμπερασμάτων της ανάλυσης. Όπως έχουμε ήδη αναφέρει, η εξαρτήσεις μεταξύ των παραγόντων της αρχιτεκτονικής αφορούν το συγκεκριμένο εξεταζόμενο σχεδιασμό και τα αποτελέσματα δε παρέχουν ενδείξεις για τη ύπαρξη σημείων αντιστάθμισης ή ευαισθησίας σε άλλες αρχιτεκτονικές. Θεωρούμε ωστόσο ότι η μέθοδος μπορεί να εφαρμοστεί σε όλα τα αρχιτεκτονικά πρότυπα όπου οι εξεταζόμενες μετρικές είναι σημαντικές. Λόγου χάρη, ο χρόνος απόκρισης μπορεί να μελετηθεί από την προτεινόμενη ανάλυση και σε άλλες αρχιτεκτονικές που δε χαρακτηρίζονται από το πρότυπο Διαχειριστή Διεργασιών [Gorton, 2006] που χαρακτηρίζει την αρχιτεκτονική επεξεργασίας συναλλαγών.

Σχετικά με την προσδοκώμενη εγκυρότητα των συμπερασμάτων, η προτεινόμενη μέθοδος έχει στέρεα θεωρητικά θεμέλια και βασίζεται σε δοκιμασμένες πρακτικές. Αναφορικά με το εξεταζόμενη αρχιτεκτονική, παροτρύνουμε τον ενδιαφερόμενο αναγνώστη να δει τους ελέγχους για την εγκυρότητα του μοντέλου, των πειραματικών παραμέτρων και των αφαιρέσεων του μοντέλου στην Ενότητα 4.2 και 4.4.5.

Άλλος ένας κίνδυνος που αναφέρεται ως δοκιμή δομικής εγκυρότητας *construct validity* στη βιβλιογραφία, αφορά το κατά πόσο οι μετρήσεις αφορούν δεδομένα σχετικά με τις υποθέσεις μας για την εξεταζόμενη αρχιτεκτονική λογισμικού. Οι υποθέσεις μας διατυπώθηκαν στην Ενότητα 4.2. Άλλες απειλές σχετικές με την εγκυρότητα δόμησης δεν αφορούν τη δική μας έρευνα διότι οι συμμετέχοντες στο πείραμα (παράγοντες και αρχιτεκτονικές επιλογές) δεν είναι υποκείμενα με υποθέσεις για το σκοπό του πειράματος (*hypothesis guessing*) ή άγχος για το αποτέλεσμά του (*evaluation apprehension*). Αξιοποιήσαμε στην ανάλυσή μας πολλές μετρικές με στόχο τη διερεύνηση της διακύμανσης της απόδοσης κατανεμημένων συναλλαγών και της διαθεσιμότητας των εξυπηρετητών με διάφορες τιμές των παραγόντων που επιδρούν στις παραπάνω ποιοτικές μετρικές. Οι ίδιες μετρικές αποδείχτηκαν αποτελε-

σματικές σε άλλες σχετικές εργασίες [Smith and Williams, 2002] που αφορούν την εκτίμηση της απόδοσης από δεδομένα προσομοίωσης. Οπότε, η απειλή στην εγκυρότητα των αποτελεσμάτων λόγω των ερευνητικών προσδοκιών μας (researcher expectancy), μπορεί να αποκλειστεί με βεβαιότητα.

Οι διαφορετικές πτυχές της αρχιτεκτονικής [Kruchten, 1995] όπως η λογική, η διαδικαστική, η αναπτυξιακή και η φυσική απεικόνιση, συνήθως αναπαριστώνται από ειδικά διαγράμματα. Η συνεπής μετατροπή των διαγραμμάτων σε μοντέλο προσομοίωσης ή πρωτότυπη υλοποίηση δε γίνεται αυτόματα και απαιτεί εμπειρία και ικανότητες. Αν οι επικρατέστερες μέθοδοι κατασκευής αρχιτεκτονικών διαγραμμάτων, όπως η UML λόγω χάρη, παρείχαν τυπική σημασιολογία για την ξεκάθαρη ερμηνεία των διαγραμμάτων, η αυτόματη μετατροπή δε θα ήταν τόσο δύσκολη υπόθεση. Η αποτελεσματικότητα της προτεινόμενης μεθόδου εξαρτάται καίρια από τη σωστή αναπαράσταση της αρχιτεκτονικής. Θεωρούμε ωστόσο ότι καθώς τα εργαλεία σχεδίασης ωριμάζουν και η σχεδιαστική ομάδα εξοικειώνεται με τις δυνατότητες και τα προβλήματα των εργαλείων, ο κίνδυνος εισαγωγής λαθών κατά τη μετατροπή σταδιακά θα εκλείψει.

Εκτιμούμε ότι η πρόοδος και η ωρίμανση των τεχνικών ανάπτυξης λογισμικού βασιζόμενες σε μοντέλα (Model-Driven Software Development [Stahl et al., 2006]) τελικά θα οδηγήσουν στην επέκταση των εργαλείων και των γλωσσών ώστε να παρέχουν τυπική σημασιολογία. Τότε, η αυτόματη ή η υποβοηθούμενη μετατροπή των διαγραμμάτων σε τυπικά μοντέλα θα επιτρέψει τον εύκολο πειραματισμό και θα διευκολύνει την εφαρμογή της προτεινόμενης μεθόδου.

## 5.5 Συμπεράσματα

Στις προηγούμενες ενότητες παρουσιάσαμε μία μέθοδο που στοχεύει στη αποκάλυψη των στατιστικά σημαντικών σημείων αντιστάθμισης ποιοτικών μετρικών αρχιτεκτονικών λογισμικού και τη συσχέτιση της παρατηρούμενης διακύμανσης των τιμών των μετρικών της

εξεταζόμενης αρχιτεκτονικής. Η ανάλυση των δύο σταδίων που περιγράψαμε είναι γενική αφού δε προϋποθέτει κάποια συγκεκριμένη μετρική ή εξειδικευμένη τεχνική απόκτησης των πειραματικών δεδομένων.

Η προτεινόμενη μέθοδος εφαρμόστηκε στην αποτίμηση δεδομένων που αφορούν την αρχιτεκτονική κατανεμημένων συναλλαγών και χαρακτηρίζεται από 16 ποιοτικές μετρικές. Τα δεδομένα αποκομίσθηκαν από το προσομοιωτή κατανεμημένων συναλλαγών ACID Sim Tools. Η ανάλυση παραγόντων του πρώτου σταδίου της μεθόδου αποκάλυψε ισχυρή συσχέτιση ορισμένων μετρικών με λανθάνοντα ποιοτικό χαρακτηριστικό που σχετίζεται με το κόστος σε προσβάσεις και καταγραφή δεδομένων σε μόνιμα μέσα αποθήκευσης και την ανάνηψη από αποτυχίες του διακομιστή. Η διαμεταγωγή των τοπικών συναλλαγών βρέθηκε αρνητικά συσχετισμένη με το παραπάνω ποιοτικό χαρακτηριστικό και δε βρέθηκε κάποια συσχέτιση μεταξύ της διαμεταγωγής τοπικών και κατανεμημένων συναλλαγών. Τα χαρακτηριστικά της συμπεριφοράς των τοπικών συναλλαγών, που αποκαλύφθηκαν από την προτεινόμενη μέθοδο, δεν ήταν αναμενόμενα και ούτε μπορούσαμε να τα μαντέψουμε εμπειρικά.

Τα θηκοδιαγράμματα των τριών συνιστωσών που βρέθηκαν από την ανάλυση παραγόντων αποκαλύπτουν σημεία ευαισθησίας και σημεία αντιστάθμισης. Αποκαλύπτουν δηλαδή ένα η παραπάνω παράγοντες που επιδρούν σε πολλές μετρικές που δεν μπορούν να βελτιστοποιηθούν ταυτόχρονα. Το δεύτερο στάδιο της μεθόδου αφορά την ανάλυση διακύμανσης για τις συνιστώσες που αποκάλυψε η ανάλυση παραγόντων. Σε αυτό το στάδιο αποδίδουμε τη παρατηρούμενη διακύμανση των τιμών των συνιστωσών σε συγκεκριμένα αρχιτεκτονικά χαρακτηριστικά και των μεταξύ τους αλληλεπιδράσεις.

Η πολυπαραγωντική ανάλυση καθώς και το ζήτημα της πρόβλεψης της ποιοτικής συμπεριφοράς ενός συστήματος λογισμικού με βάση μετρικές που δεν εξετάστηκαν στην παρούσα έρευνα [Katsaros et al., 2006], αποτελούν ενδιαφέρουσες προοπτικές και επεκτάσεις στην προτεινόμενη μέθοδο.

## Κεφάλαιο 6

### Επίλογος

Τα συστήματα επεξεργασίας κατανεμημένων συναλλαγών είναι σημαντικά για πολλούς λόγους. Είναι σημαντικά για τις εφαρμογές που βρίσκουν οι κατανεμημένες συναλλαγές στη βιομηχανία, στο τραπεζικό σύστημα και οπουδήποτε αλλού απαιτούνται οι ACID εγγυήσεις που παρέχουν. Τα διαφορετικά περιβάλλοντα εκτέλεσης καθώς και οι διαφορετικές απαιτήσεις που τίθενται από τις ποικίλες εφαρμογές οδήγησαν στην εκτεταμένη διερεύνηση των αρχιτεκτονικών επιλογών των κατανεμημένων συναλλαγών. Έχουν προταθεί πολλές παραλλαγές του κλασσικού πρωτοκόλλου που αποσκοπούν στο να ικανοποιήσουν τις απαιτήσεις και τους περιορισμούς που επιβάλλουν οι διαφορετικές εφαρμογές και τα περιβάλλοντα εκτέλεσης. Από σχεδιαστικής άποψης, οι προτεινόμενες αλλαγές του βασικού πρωτοκόλλου ενδέχεται να έχουν απροσδόκητες αρνητικές επιπτώσεις στην απόδοση και τη διαθεσιμότητα επεξεργασίας συναλλαγών και ενδέχεται να ακυρώνουν κάποιες από τις ιδιότητες που ικανοποιεί το κλασσικό πρωτόκολλο συναλλαγών.

Η εκτίμηση των επιπτώσεων των σχεδιαστικών επιλογών σε επίπεδο αρχιτεκτονικής συστήματος μπορεί να βασιστεί στην ανάλυση δεδομένων που συλλέγονται με προσομοίωση των εν λόγω συστημάτων. Ωστόσο, η πολυπλοκότητα που απορρέει κατά κύριο λόγο από την ασύγχρονη ανταλλαγή μηνυμάτων μεταξύ των προσομοιωμένων οντοτήτων, κάνει πολύ πιθανή την εισαγωγή λαθών στην υλοποίηση, που αναπόφευκτα οδηγεί σε λαν-



θασμένα συμπεράσματα. Τέτοια λάθη μπορούν να αποφευχθούν με την απαλοιφή του ανθρώπινου παράγοντα μέσω της αυτόματης σύνθεσης μέρους της υλοποίησης από έγκυρα μοντέλα συναλλαγών. Η αυτόματη σύνθεση κώδικα μπορεί να χρησιμοποιηθεί όχι μόνο για τη παραγωγή προσομοίωσης, αλλά και στην υλοποίηση του συστήματος επεξεργασίας συναλλαγών.

Επιπλέον, στα προηγμένα πρωτόκολλα συναλλαγών, με τη κατάργηση ή τη μη αυστηρή εφαρμογή ορισμένων ACID εγγυήσεων του κλασσικού πρωτοκόλλου επεξεργασίας συναλλαγών, δεν είναι πλέον ξεκάθαρο ποιες από τις ιδιότητες του κλασσικού πρωτοκόλλου κληροδοτούνται στην προηγμένη εκδοχή του. Στην παρούσα εργασία προτείνουμε την μέθοδο ελέγχου μοντέλων για την επικύρωση των ιδιοτήτων ορθότητας των προηγμένων μοντέλων συναλλαγών.

## 6.1 Σύνοψη της ερευνητικής συνεισφοράς

Για την τυπική ανάλυση συστημάτων επεξεργασίας συναλλαγών απαιτείται η κατάλληλη αναπαράσταση του συστήματος από ένα μοντέλο που αν και αφαιρετικό, διατηρεί όλες τις απαραίτητες πληροφορίες που αφορούν τις ιδιότητες των συναλλαγών. Στο Κεφάλαιο 3 δείξαμε ένα τρόπο αναπαράστασης του συστήματος μέσω αυτομάτων και των μεταξύ τους αλληλεπιδράσεων. Οι ιδιότητες ορθότητας ελέγχονται με βάση ένα μοντέλο που περιγράφεται από μη αιτιοκρατικά αυτόματα.

Η επιλεγμένη αναπαράσταση του συστήματος και ο αλγόριθμος που αναπτύξαμε για τον έλεγχο των ιδιοτήτων ορθότητας, επιτρέπουν την έκφραση και την επικύρωση τριών ειδών ιδιοτήτων: προσεγγισιμότητας, ασφάλειας και βιωσιμότητας. Διαπιστώσαμε την αδυναμία ελέγχου ιδιοτήτων σε περιπτώσεις που ο χώρος καταστάσεων είναι ιδιαίτερα μεγάλος και απαιτείται πολλός χρόνος για τον έλεγχο όλων των δυνατών εκτελέσεων που αφορούν την εξεταζόμενη ιδιότητα. Εισάγουμε μία διεργασία που μπορεί να περιορίσει το χώρο των καταστάσεων και να μειώσει σε αποδεκτά πλαίσια τον απαιτούμενο χρόνο ελέγχου.

Εξετάζουμε την αποτελεσματικότητα της μεθόδου ελέγχου μοντέλων στο βασικό μοντέλο συναλλαγών και σε δύο παραλλαγές του. Επιπλέον, εξετάζουμε ένα παράδειγμα προηγμένου πρωτοκόλλου συναλλαγών, το ενθουλακόμενο μοντέλο, και δύο παραλλαγές του. Τα αποτελέσματα τις έρευνας παρουσιάστηκαν στις εργασίες [Mentis and Katsaros, 2009, Mentis and Katsaros, 2012a].

Στο Κεφάλαιο 4 αναλύουμε την απόδοση και τη διαθεσιμότητα επεξεργασίας συστημάτων με ασύγχρονη επικοινωνία. Εξετάζουμε συστήματα επεξεργασίας κατανεμημένων συναλλαγών που ανήκουν στην ευρύτερη κατηγορία των συστημάτων επεξεργασίας ασύγχρονων μηνυμάτων. Με αφετηρία ένα έγκυρο μοντέλο του συστήματος, προτείνουμε μία μέθοδο για την ανάλυση της απόδοσης και της διαθεσιμότητας επεξεργασίας με προσομοίωση του συστήματος που μας ενδιαφέρει.

Η συμπεριφορά του συστήματος εξετάζεται με προσομοίωση σε διάφορα βασικά σενάρια εκτέλεσης. Τα σενάρια αφορούν περιπτώσεις χρήσης που εκτελούνται συχνά και διαφοροποιούνται στις απαιτήσεις που έχουν σε πόρους λογισμικού, μνήμη και επεξεργαστική ισχύ. Η δεύτερη κατηγορία σεναρίων εκτέλεσης αφορά περιπτώσεις που ενώ εμφανίζονται σπάνια κατά την λειτουργία του συστήματος, επιβαρύνουν με ακραίο τρόπο την απόδοση και την διαθεσιμότητα επεξεργασίας. Η μέθοδος ελέγχει αν ικανοποιούνται οι στόχοι απόδοσης στα εξεταζόμενα σενάρια. Τα βήματα της προτεινόμενης μεθόδου εφαρμόζονται επαναληπτικά μέχρι την επίτευξη των στόχων. Σε κάθε μεταβολή του μοντέλου και επακόλουθη επανάληψη των βημάτων της μεθόδου, επαληθεύεται η ορθότητα ως προς τις ιδιότητες που πρέπει να ικανοποιεί. Όσον αφορά το μοντέλο προσομοίωσης, ελέγχεται η εγκυρότητα των αφαιρέσεων του μοντέλου και των υποθέσεων που αφορούν τις τιμές των παραμέτρων προσομοίωσης. Τα αποτελέσματα της εφαρμογής της μεθόδου αναλύονται στην εργασία [Mentis and Katsaros, 2012b] που έχει υποβληθεί για δημοσίευση. Η μέθοδος εφαρμόστηκε σε αποτελέσματα προσομοίωσης που αποκομίσθηκαν από τον προσομοιωτή ACID Sim Tools [Mentis et al., 2008].

Σε αρχιτεκτονικές συστημάτων λογισμικού με περίπλοκες αλληλεπιδράσεις παρατηρούνται συχνά αλληλεπιδράσεις μεταξύ των παραγόντων που επηρεάζουν ποιοτικές μετρικές,

όπως την απόδοση και την διαθεσιμότητα επεξεργασίας. Οι αλληλεπιδράσεις μεταξύ των σχεδιαστικών παραγόντων, αν και επιδρούν στην ποιότητα της αρχιτεκτονικής, δεν μπορούν να ποσοτικοποιηθούν εύκολα και γιαυτό τις αποκαλούμε λανθάνοντες παράγοντες ποιότητας. Στο Κεφάλαιο 5 παρουσιάζουμε μία μέθοδο για την αποτίμηση των επιπτώσεων παραγόντων με σύνθετες αλληλεπιδράσεις στην ποιότητα επεξεργασίας. Η μέθοδος εφαρμόστηκε στην ανάλυση δεδομένων που αφορούν αρχιτεκτονικές κατανεμημένων συναλλαγών, αλλά είναι γενική από την άποψη ότι δεν προϋποθέτει την ύπαρξη κάποιας συγκεκριμένης μετρικής ούτε εξαρτάται από κάποιο τρόπο συλλογής των πειραματικών δεδομένων (προσομοίωση, υλοποίηση πρωτοτύπου).

Από την εφαρμογή της μεθόδου σε δεδομένα προσομοίωσης αρχιτεκτονικών κατανεμημένων συναλλαγών, οι αρχικά 16 μετρικές ποιότητας περιορίστηκαν μόνο σε τρεις που βρέθηκαν να είναι μεταξύ τους ασυσχέτιστες. Η ανάλυση παραγόντων του πρώτου σταδίου της μεθόδου αποκάλυψε ισχυρή συσχέτιση ορισμένων μετρικών με λανθάνον ποιοτικό χαρακτηριστικό που σχετίζεται με το κόστος σε προσβάσεις και καταγραφή δεδομένων σε μέσα μόνιμης αποθήκευσης και την ανάνηψη από αποτυχίες πτώσης συστήματος. Η διαμεταγωγή των τοπικών συναλλαγών βρέθηκε αρνητικά συσχετισμένη με το παραπάνω ποιοτικό χαρακτηριστικό ενώ δε βρέθηκε συσχέτιση μεταξύ της διαμεταγωγής τοπικών και κατανεμημένων συναλλαγών. Τα χαρακτηριστικά της συμπεριφοράς των τοπικών συναλλαγών, που αποκαλύφθηκαν από την προτεινόμενη μέθοδο, δεν ήταν αναμενόμενα και δε θα μπορούσαν να προβλεφθούν εμπειρικά. Τα αποτελέσματα και η περιγραφή της προτεινόμενης μεθόδου παρουσιάζονται αναλυτικά στις εργασίες [Mentis et al., 2009, Mentis et al., 2010].

## 6.2 Ερευνητικές προοπτικές

Οι διαφορετικές πτυχές ενός μοντέλου συναλλαγών, τα αυτόματα, οι ενέργειες που καλούνται στις μεταβάσεις τους και τα μηνύματα που στέλνονται ή ακυρώνονται από τις καλούμενες ενέργειες, περιγράφονται σε αρχεία κειμένου με απλή και κατανοητή δομή. Ωστόσο, η

διαδικασία ορισμού της αρχιτεκτονικής και η επαλήθευση των ιδιοτήτων ορθότητας θα μπορούσε να διευκολυνθεί από γραφικά σχεδιαστικά εργαλεία ορισμού αρχιτεκτονικής. Υπάρχουν είδη πολλά και αξιόλογα εργαλεία σχεδιασμού με την γλώσσα UML, τόσο σε πλατφόρμα ανοιχτού λογισμικού όσο και σε εμπορικά προϊόντα. Τα διαγράμματα καταστάσεων (statechart), είναι τα καταλληλότερα για την περιγραφή των αυτομάτων του μοντέλου. Η υλοποίηση της περιγραφής μοντέλων στην μορφή XMI, που υποστηρίζεται από τα περισσότερα εργαλεία γραφικής επεξεργασίας UML διαγραμμάτων, θα επιτρέψει τον καθορισμό του μοντέλου με γραφικό τρόπο.

Στην παρούσα μορφή του εργαλείου ACID Model Checker, οι ιδιότητες ορθότητας διατυπώνονται από συναρτήσεις διατυπωμένες στη γλώσσα Haskell [Peyton Jones, 2003]. Σκοπεύουμε να επεκτείνουμε τις δυνατότητες του εργαλείου, ώστε να υποστηρίξει διαφορετικό φορμαλισμό για την διατύπωση των ιδιοτήτων ορθότητας και να μην απαιτείται γνώση της γλώσσας Haskell για τη χρήση του ελεγκτή μοντέλου.

Η προεπεξεργασία του μοντέλου αφαιρεί, υπό προϋποθέσεις, μεταβάσεις που δεν παράγουν ούτε ακυρώνουν μηνύματα βελτιώνοντας έτσι το χρόνο για τον έλεγχο ιδιοτήτων ορθότητας. Σε σύνθετα μοντέλα συναλλαγών, όπως το ενθυλακόμενο μοντέλο, τα περιθώρια βελτίωσης από την προεπεξεργασία είναι σημαντικά. Υπάρχουν ωστόσο και άλλες δυνατότητες βελτίωσης του αλγορίθμου ελέγχου ιδιοτήτων ορθότητας. Με την ενσωμάτωση τεχνικών μείωσης του χώρου καταστάσεων με μερική διάταξη (partial order reduction, θα μπορούσε να βελτιωθεί περαιτέρω ο χρόνος ελέγχου της ορθότητας των μοντέλων συναλλαγών. Σε ορισμένες περιπτώσεις, η σειρά εκτέλεσης των πιθανών μεταβάσεων δεν έχει σημασία για την ορθότητα κάποιας ιδιότητας. Η τεχνική μερικής διάταξης εκμεταλλεύεται την παραπάνω ιδιότητα του χώρου καταστάσεων για τον περιορισμό του.

Άλλη τεχνική που θα μπορούσε να βελτιώσει το χρόνο ελέγχου προέρχεται από την κατηγορία των  $A^*$  αλγορίθμων. Πρόκειται για κατευθυνόμενη διερεύνηση του χώρου καταστάσεων με στόχο το γρήγορο εντοπισμό καταστάσεων που παραβιάζουν κάποια ιδιότητα ορθότητας. Σε αυτούς τους αλγορίθμους [Yang and Dill, 1998, Dräger et al., 2006], η εξερεύνηση του χώρου καταστάσεων καθοδηγείται από μία εκτίμηση της απόστασης της

τρέχουσας κατάστασης από την πλησιέστερη κατάσταση όπου παραβιάζεται η ιδιότητα ορθότητας. Η τεχνική μπορεί να εντοπίσει ταχύτερα τις καταστάσεις όπου παραβιάζονται ιδιότητες αλλά απαιτείται εξαντλητική εξερεύνηση του χώρου καταστάσεων αν η ιδιότητα δεν παραβιάζεται.

Μία εναλλακτική προσέγγιση τυπικής ανάλυσης μοντέλων συναλλαγών με σημαντικές προοπτικές θα μπορούσε να είναι η υιοθέτηση ενός πλαισίου μοντελοποίησης με αρθρώματα που ορίζουν συμπεριφορά (π.χ αυτόματα) έτσι ώστε οι μεταξύ τους αλληλεπιδράσεις να περιγράφονται χωρίς να εισάγεται στο σύστημα επιπλέον συμπεριφορά. Αυτό σημαίνει ότι με τις αλληλεπιδράσεις των αρθρωμάτων περιορίζεται η υπάρχουσα συμπεριφορά τους ώστε να ανταποκρίνεται στις προδιαγραφές του συστήματος που εξετάζεται. Ένα τέτοιο πλαίσιο ανάλυσης είναι το BIP (Behaviour, Interactions, Priorities)[Göblier and Sifakis, 2005] που υποστηρίζει προηγμένες τεχνικές επαλήθευσης με στατική ανάλυση της δομής του μοντέλου. Αυτές οι τεχνικές εφαρμόζονται αυξητικά κατά τη σχεδίαση και όχι σε μοντέλο που παράγεται μετά την ολοκλήρωση της σχεδίασης. Η προσέγγιση αποδίδει δομικά ορθή σχεδίαση (correct by construction) και υποστηρίζει όλα τα στάδια ανάπτυξης ενός συστήματος [Basu et al., 2011] ξεκινώντας από τη σχεδίαση της εφαρμογής μέχρι και την ανάλυση της αρχιτεκτονικής του συστήματος και την αυτόματη σύνθεση κώδικα κατανεμημένης υλοποίησης. Αν και για την ώρα υποστηρίζεται μόνο η ανάλυση ιδιοτήτων ασφάλειας και η αποφυγή αδιεξόδων, το σημαντικό πλεονέκτημα είναι η δυνατότητα κλιμάκωσης της ανάλυσης σε μοντέλα με πολύ μεγάλο χώρο καταστάσεων σε αντίθεση με τη τεχνική ελέγχου μοντέλων. Πρόσφατα εφαρμόσαμε την παραπάνω μέθοδο στην τυπική ανάλυση της σύνθεσης υπηρεσιών ιστού με τη γλώσσα BPEL [Stachtari et al., 2012].

Έχουμε αναφερθεί σε κάποιες από τις εφαρμογές των κατανεμημένων συναλλαγών και τις παραλλαγές στο σχεδιασμό πρωτοκόλλων συναλλαγών με στόχο την εξυπηρέτηση των ειδικών αναγκών που προβάλλουν στα διαφορετικά περιβάλλοντα εκτέλεσης συναλλαγών. Τα τελευταία χρόνια, οι κατανεμημένες συναλλαγές προτάθηκαν ως εναλλακτικός τρόπος συντονισμού και συνεργασίας διαδικασιών που εκτελούνται σε κάποιο από τους πολλούς διαφορετικούς πυρήνες των σύγχρονων υπολογιστικών συστημάτων. Οι μνήμες συναλλα-

γών (βλέπε [Marathe and Scott, 2004] για μία επισκόπηση της ερευνητικής περιοχής) διαφοροποιούνται σε μεγάλο βαθμό από το κλασικό πρωτόκολλο συναλλαγών. Καταρχάς, οι διεργασίες δεν επικοινωνούν μέσω δικτύου οπότε όλα τα ζητήματα που σχετίζονται με την προσωρινή απώλεια επικοινωνίας ή τη ταχύτητα μετάδοσης μηνυμάτων δεν αφορούν τις μνήμες συναλλαγών. Επιπλέον, δεν απαιτείται από τις μνήμες συναλλαγών να διαθέτουν ανοχή σε αποτυχίες πτώσης του συστήματος. Αναμένεται ότι οι εφαρμογές με παράλληλη επεξεργασία θα είναι το κυρίαρχο μοντέλο υπολογισμού και οι μνήμες συναλλαγών έχουν πολλά να προσφέρουν προς αυτή την κατεύθυνση. Αντίθετα από τις κλασικές διαδικασίες συγχρονισμού διεργασιών όπως κλείδωμα κοινόχρηστων μεταβλητών, σηματοφόροι και κοινόχρηστη μνήμη, οι συναλλαγές μνήμης επιτρέπουν τη σύνθεση μικρών και εύκολα διαχειρίσιμων προγραμμάτων σε μεγαλύτερα και πιο σύνθετα με καλώς ορισμένο τρόπο [Harris et al., 2008]. Πιστεύουμε ότι η εφαρμογή των μεθόδων που προτείνουμε στην παρούσα διατριβή πάνω στις μνήμες συναλλαγών θα αποδώσει ενδιαφέροντα αποτελέσματα.

# Παράρτημα I

## Παράρτημα

### I.1 Χρήση του εργαλείου ACID Sim Tools στα πλαίσια της διαδικασίας αποτίμησης αρχιτεκτονικής συναλλαγών

Τα απαραίτητα βήματα για την αποτελεσματική εφαρμογή της μεθόδου αποτίμησης αρχιτεκτονικών επιλογών του Κεφαλαίου 5 σε συνδυασμό με το εργαλείο ACID Sim Tools είναι τα ακόλουθα:

1. Ορίζουμε τις διάφορες πτυχές της αρχιτεκτονικής [Kruchten, 1995] σε UML ή κάποια άλλη γλώσσα ορισμού αρχιτεκτονικών.
2. Εξετάζουμε αν τα αρθρώματα που παρέχει το ACID Sim Tools μπορούν να προσημειώσουν όλα τα επιθυμητά αρχιτεκτονικά χαρακτηριστικά. Αν δεν αρκούν, αξιοποιούμε τη γλώσσα ορισμού του ACID Sim Tools για την υλοποίηση των υπόλοιπων χαρακτηριστικών (βλέπε Κεφάλαιο 3).
3. Μετατρέπουμε τον ορισμό της αρχιτεκτονικής σε μοντέλο προσομοίωσης του OMNeT++

[Varga and Hornig, 2008].

4. Επιβεβαιώνουμε ότι ισχύουν οι προσδοκώμενες ιδιότητες ορθότητας του μοντέλου (βλέπε Κεφάλαιο 3).
5. Θέτουμε τις τιμές των παραμέτρων του μοντέλου προσομοίωσης όπως προκύπτουν από τον υπολογισμό των απαραίτητων πόρων. Ενδεικτικές τιμές για τις απαιτήσεις σε πόρους μπορεί να προέρχονται από μετρήσεις πρωτοτύπων υλοποιήσεων της αρχιτεκτονικής ή από υλοποιήσεις παρόμοιων αρχιτεκτονικών. Οι τιμές των παραμέτρων πρέπει να αντιστοιχούν σε πραγματικές μετρήσεις.
6. Επικύρωση του μοντέλου αναφορικά με τις απειλές στην εγκυρότητα που διατυπώθηκαν στην Ενότητα εγκυρότητα φάσης.
7. Ορίζουμε το πειραματικό σχεδιασμό.
8. Συλλέγουμε τα δεδομένα των πειραμάτων και προχωρούμε στην εφαρμογή της ανάλυσης δύο σταδίων, όπως περιγράφεται στην Ενότητα 5.2.

Τα βήματα 5 – 8 μπορεί να εκτελεστούν για σύνολα παραμέτρων που αντιστοιχούν σε διαφορετικούς φόρτους εργασίας.



# Bibliography

- [Agrawal et al., 1996] Agrawal, R., Carey, M. J., and Livny, M. (1996). Concurrency control performance modeling: Alternatives and implications. In *Performance of Concurrency Control Mechanisms in Centralized Database Systems*, pages 58–105.
- [Anderson and Rubin, 1956] Anderson, T. W. and Rubin, H. (1956). *Statistical Inference in Factor Analysis*. Cowles Foundation paper. Cowles Foundation for Research in Economics at Yale University.
- [Atluri et al., 1997] Atluri, V., Bertino, E., and Jajodia, S. (1997). A theoretical formulation for degrees of isolation in databases. *Information & Software Technology*, 39(1):47–53.
- [Baier and Katoen, 2008] Baier, C. and Katoen, J.-P. (2008). *Principles of model checking*. MIT Press.
- [Balsamo et al., 2004a] Balsamo, S., Marco, A. D., Inverardi, P., and Simeoni, M. (2004a). Model-based performance prediction in software development: A survey. *IEEE Trans. Software Eng.*, 30(5):295–310.
- [Balsamo and Marzolla, 2003] Balsamo, S. and Marzolla, M. (2003). Simulation modeling of UML software architectures.
- [Balsamo et al., 2004b] Balsamo, S., Marzolla, M., Marco, A. D., and Inverardi, P. (2004b). Experimenting different software architectures performance techniques: a case study. In Dujmovic, J. J., Almeida, V. A. F., and Lea, D., editors, *Proceedings of the Fourth*

- International Workshop on Software and Performance, WOSP 2004, Redwood Shores, California, USA, January 14-16, 2004*, pages 115–119. ACM.
- [Bardram et al., 2005] Bardram, J. E., Christensen, H. B., Corry, A. V., Hansen, K. M., and Ingstrup, M. (2005). Exploring quality attributes using architectural prototyping. In *QoSA/SOQUA*, pages 155–170.
- [Bass et al., 2003] Bass, L., Clements, P., and Kazman, R. (2003). *Software Architecture in Practice (2nd Edition)*. Addison Wesley.
- [Basu et al., 2011] Basu, A., Bensalem, S., Bozga, M., Combaz, J., Jaber, M., Nguyen, T.-H., and Sifakis, J. (2011). Rigorous component-based system design using the bip framework. *IEEE Software*, 28(3):41–48.
- [Becker et al., 2008] Becker, S., Dencker, T., and Happe, J. (2008). Model-driven generation of performance prototypes. In *SIPeW*, pages 79–98.
- [Bengtsson and Bosch, 1998] Bengtsson, P. and Bosch, J. (1998). Scenario-based software architecture reengineering. In *Proceedings of the 5th International Conference on Software Reuse, ICSR '98, Washington, DC, USA*. IEEE Computer Society.
- [Berander et al., 2006] Berander, P., Damm, L. O., Eriksson, J., Gorschek, T., Henningsson, K., Jönsson, P., Kaagström, S., Milicic, D., Maartensson, F., Rönkkö, K., and Tomaszewski, P. (2006). *Software quality attributes and trade-offs*. Blekinge Institute of Technology.
- [Bérard et al., 2001] Bérard, B., Bidoit, M., Finkel, A., Laroussinie, F., Petit, A., Petrucci, L., and Schnoebelen, Ph. (2001). *Systems and Software Verification. Model-Checking Techniques and Tools*. Springer.
- [Bernstein et al., 1987] Bernstein, P. A., Hadzilacos, V., and Goodman, N. (1987). *Concurrency Control and Recovery in Database Systems*. Addison-Wesley.

- [Bernstein and Newcomer, 2009] Bernstein, P. A. and Newcomer, E. (2009). *Principles of Transaction Processing*. Morgan Kaufmann Series in Data Management Systems. Morgan Kaufmann Publishers.
- [Blom, 1958] Blom, G. (1958). *Statistical Estimates and Transformed Beta Variables*. Wiley, New York.
- [Chrysanthis and Ramamritham, 1990] Chrysanthis, P. K. and Ramamritham, K. (1990). ACTA. A framework for specifying and reasoning about transaction structure and behavior. *SIGMOD Record (ACM Special Interest Group on Management of Data)*, 19(2):194–203.
- [Chrysanthis and Ramamritham, 1994] Chrysanthis, P. K. and Ramamritham, K. (1994). Synthesis of extended transaction models using ACTA. *ACM Transactions on Database Systems*, 19(3):450–491.
- [Chrysanthis et al., 1998] Chrysanthis, P. K., Samaras, G., and Al-Houmailly, Y. J. (1998). *Recovery and Performance of Atomic Commit Processing in Distributed Database Systems*, pages 370–416. *Recovery Mechanisms in Database Systems*. Printice Hall.
- [Clarke et al., 1999] Clarke, E., Grunberg, O., and Peled, E. (1999). *Model Checking*. MIT Press.
- [Clements et al., 2001] Clements, P., Kazman, R., and Klein, M. (2001). *Evaluating Software Architectures: Methods and Case Studies*. Addison-Wesley.
- [Crnkovic et al., 2004] Crnkovic, I., Larsson, M., and Preiss, O. (2004). Concerning predictability in dependable component-based systems: Classification of quality attributes. In de Lemos, R., Gacek, C., and Romanovsky, A. B., editors, *WADS*, volume 3549 of *Lecture Notes in Computer Science*, pages 257–278. Springer.
- [Dobrica and Niemelä, 2002] Dobrica, L. and Niemelä, E. (2002). A survey on software architecture analysis methods. *IEEE Trans. Software Eng*, 28(7):638–653.

- [Dräger et al., 2006] Dräger, K., Finkbeiner, B., and Podelski, A. (2006). Directed model checking with distance-preserving abstractions. In Valmari, A., editor, *Model Checking Software*, volume 3925 of *Lecture Notes in Computer Science*, pages 19–34. Springer Berlin / Heidelberg.
- [Eguiluz and Barbacci, 2003] Eguiluz, H. and Barbacci, M. R. (2003). *Interactions Among Techniques Addressing Quality Attributes*. Technical report. Carnegie Mellon University, Software Engineering Institute.
- [Elmagarmid, 1992] Elmagarmid, A. K., editor (1992). *Database Transaction Models for Advanced Applications*. Morgan Kaufmann.
- [Field, 2009] Field, A. P. (2009). *Discovering Statistics Using SPSS: (And Sex and Drugs and Rock 'n' Roll)*. Introducing Statistical Methods Series. Sage.
- [Garcia-Molina and Salem, 1987] Garcia-Molina, H. and Salem, K. (1987). Sagas. In Dayal, U. and Traiger, I. L., editors, *SIGMOD Conference*, pages 249–259. ACM Press.
- [Gorton, 2006] Gorton, I. (2006). *Essential Software Architecture*. Springer-Verlag New York, Inc., Secaucus, NJ, USA.
- [Gößler and Sifakis, 2005] Gößler, G. and Sifakis, J. (2005). Composition for component-based modeling. *Sci. Comput. Program.*, 55(1-3):161–183.
- [Grefen et al., 2001] Grefen, P. W. P. J., Vonk, J., and Apers, P. M. G. (2001). Global transaction support for workflow management systems: from formal specification to practical implementation. *VLDB J.*, 10(4):316–333.
- [Grunske, 2006] Grunske, L. (2006). Identifying "good" architectural design alternatives with multi-objective optimization strategies. In Osterweil, L. J., Rombach, H. D., and Sofa, M. L., editors, *ICSE*, pages 849–852. ACM.

- [Grunske, 2007] Grunske, L. (2007). Early quality prediction of component-based systems - A generic framework. *Journal of Systems and Software*, 80(5):678–686.
- [Gupta et al., 1996] Gupta, R., Haritsa, J., Ramamritham, K., and Seshadri, S. (1996). Commit processing in distributed real-time database systems. In *Proceedings of 17th IEEE Real-Time Systems Symposium*.
- [Harris, 2001] Harris, R. J. (2001). *A Primer of Multivariate Statistics*. Lawrence Erlbaum Associates.
- [Harris et al., 2008] Harris, T., Marlow, S., Jones, S. P., and Herlihy, M. (2008). Composable memory transactions. *Commun. ACM*, 51:91–100.
- [Herlihy and Moss, 1993] Herlihy, M. and Moss, J. E. B. (1993). Transactional memory: architectural support for lock-free data structures. In *Proceedings of the 20th annual international symposium on computer architecture, ISCA '93*, pages 289–300, New York, NY, USA. ACM.
- [Holzmann, 2003] Holzmann, G. J. (2003). *The Spin Model Checker, Primer and Reference Manual*. Addison-Wesley, Reading, Massachusetts.
- [Johnson et al., 2007] Johnson, J. E., Langworthy, D. E., Lamport, L., and Vogt, F. H. (2007). Formal specification of a web services protocol. *J. Log. Algebr. Program*, 70(1):34–52.
- [Katsaros et al., 2001] Katsaros, P., Angelis, L., and Lazos, C. (2001). Applied multire-sponse metamodeling for queuing network simulation experiments: Problems and perspectives. In *Proc. of the 4th EUROSIM Congress on Modelling and Simulation, EUROSIM, Delfts, The Netherlands*.
- [Katsaros et al., 2007] Katsaros, P., Angelis, L., and Lazos, C. (2007). Performance and effectiveness trade-off for checkpointing in fault-tolerant distributed systems. *Concurrency and Computation: Practice and Experience*, 19(1):37–63.

- [Katsaros et al., 2006] Katsaros, P., Iakovidou, N., and Soldatos, T. (2006). Evaluation of composite object replication schemes for dependable server applications. *Information & Software Technology*, 48(9):795–806.
- [Kazman, 1994] Kazman, R. (1994). SAAM: A method for analyzing the properties of software architectures. In *in Proceedings of the 16th International Conference on Software Engineering*, pages 81–90.
- [Kelton and Law, 1984] Kelton, W. D. and Law, A. M. (1984). An analytical evaluation of alternative strategies in steady-state simulation. *Operations Research*, 32(1):169–184.
- [Kitchenham et al., 2002] Kitchenham, B., Pfleeger, S. L., Pickard, L., Jones, P., Hoaglin, D. C., Emam, K. E., and Rosenberg, J. (2002). Preliminary guidelines for empirical research in software engineering. *IEEE Trans. Software Eng.*, 28(8):721–734.
- [Klein and Weiss, 2009] Klein, J. and Weiss, D. (2009). *What is architecture?* O'Reilly Media.
- [Kruchten, 1995] Kruchten, P. (1995). Architecture blueprints - the "4+1" view model of software architecture. In *TRI-Ada Tutorials*, pages 540–555.
- [Krzanowski, 1993] Krzanowski, W. (1993). *Principles of Multivariate Analysis: A User's Perspective*. Oxford University Press.
- [Lampson and Sturgis, 1979] Lampson, B. W. and Sturgis, H. E. (1979). Crash recovery in a distributed data storage system.
- [Litoiu et al., 2000] Litoiu, M., Rolia, J. A., and Serazzi, G. (2000). Designing process replication and activation: A quantitative approach. *IEEE Trans. Software Eng.*, 26(12):1168–1178.

- [Liu et al., 1998] Liu, M. L., Agrawal, D., and Abbadi, A. E. (1998). The performance of two phase commit protocols in the presence of site failures. *Distrib. Parallel Databases*, 6(2):157–182.
- [Losavio, 2002] Losavio, F. (2002). Quality models to design software architecture. *Journal of Object Technology*, 1(4):165–178.
- [Lynch and Merritt, 1986] Lynch and Merritt (1986). Introduction to the theory of nested transactions. In *ICDT: International Conference on Database Theory*.
- [Marathe and Scott, 2004] Marathe, V. J. and Scott, M. L. (2004). A qualitative survey of modern software transactional memory systems. Technical Report TR 839, University of Rochester Computer Science Dept.
- [Martin and Ramamritham, 1997] Martin, C. P. and Ramamritham, K. (1997). Toward formalizing recovery of (advanced) transactions. In *Advanced Transaction Models and Architectures*.
- [Martis, 2006] Martis, M. S. (2006). Validation of simulation based models: A theoretical outlook.
- [Mentis and Katsaros, 2009] Mentis, A. and Katsaros, P. (2009). The ACID model checker and code generator for transaction processing. In *Proceedings of the High Performance Computing & Simulation Conference*, Leipzig, Germany.
- [Mentis and Katsaros, 2012a] Mentis, A. and Katsaros, P. (2012a). Model checking and code generation for transaction processing software. *Concurrency and Computation: Practice and Experience*, 24(7):711–722.
- [Mentis and Katsaros, 2012b] Mentis, A. and Katsaros, P. (2012b). Simulation-based performance evaluation of systems with asynchronous event processing: an application on transaction models. *Simulation Modelling Practice and Theory*. To appear.

- [Mentis et al., 2008] Mentis, A., Katsaros, P., and Angelis, L. (2008). ACID sim tools: A simulation framework for distributed transaction processing architectures. In *Processing of the 1st International Conference on Simulation Tools and Techniques for Communications, Networks and Systems (SimulationWorks Industry Track)*, Marseille, France.
- [Mentis et al., 2009] Mentis, A., Katsaros, P., and Angelis, L. (2009). Synthetic metrics for evaluating performance of software architectures with complex tradeoffs. In *Proceedings of the 35th EUROMICRO Conference on Software Engineering and Advanced Applications (SEAA)*, pages 237–242, Patra, Greece. IEEE Computer Society.
- [Mentis et al., 2010] Mentis, A., Katsaros, P., Angelis, L., and Kakarontzas, G. (2010). Quantification of interacting runtime qualities in software architectures: insights from transaction processing in client-server architectures. *Information and Software Technology*, 52(12):1331–1345.
- [Mohan et al., 1986] Mohan, C., Lindsay, B., and Obermarck, R. (1986). Transaction management in the R\* distributed database management system. *ACM Trans. Database Syst.*, 11(4):378–396.
- [Moss, 1982] Moss, J. E. B. (1982). Nested transactions and reliable distributed computing. In *Proc. 2nd Symp. on Reliability in Distributed Software and Database Systems*, page 33.
- [Nouali-Taboudjemat et al., 2010] Nouali-Taboudjemat, N., Chehbour, F., and Drias, H. (2010). On performance evaluation and design of atomic commit protocols for mobile transactions. *Distributed and Parallel Databases*, 27:53–94.
- [Olumofin and Mistic, 2007] Olumofin, F. G. and Mistic, V. B. (2007). A holistic architecture assessment method for software product lines. *Information & Software Technology*, 49(4):309–323.
- [OMG, 2003] OMG (2003). Transaction service specification (version 1.4).



- [Paul et al., 2003] Paul, H. S., Gupta, A., and Badrinath, R. (2003). Performance comparison of checkpoint and recovery protocols. *Concurrency and Computation: Practice and Experience*, 15(15):1363–1386.
- [Peyton Jones, 2003] Peyton Jones, S. (2003). Special issue: Haskell 98 language and libraries. *Journal of Functional Programming*, 13.
- [Reussner and Firus, 2005] Reussner, R. and Firus, V. (2005). Introduction to overlapping attributes. In Eusgeld, I., Freiling, F. C., and Reussner, R., editors, *Dependability Metrics*, volume 4909 of *Lecture Notes in Computer Science*, pages 243–244. Springer.
- [Samaras et al., 1993] Samaras, G., Britton, K., Citron, A., and Mohan, C. (1993). Two-phase commit optimizations and tradeoffs in the commercial environment. In *Data Engineering, 1993. Proceedings. Ninth International Conference on*, pages 520–529.
- [Sargent, 2005] Sargent, R. G. (2005). Verification and validation of simulation models. In *Proceedings of the 37th conference on Winter simulation, WSC '05*, pages 130–143. Winter Simulation Conference.
- [Serrano-Alvarado et al., 2004] Serrano-Alvarado, P., Roncancio, C., and Adiba, M. (2004). A survey of mobile transactions.
- [Shaw and Garlan, 1996] Shaw, M. and Garlan, D. (1996). *Software architecture - perspectives on an emerging discipline*. Prentice Hall.
- [Smith and Williams, 2002] Smith, C. U. and Williams, L. G. (2002). *Performance Solutions: a practical guide to creating responsive, scalable software*. Addison-Wesley.
- [Soley and Stone, 1995] Soley, R. M. and Stone, C. M. (1995). *Object management architecture guide (3rd ed.)*. John Wiley & Sons, Inc., New York, NY, USA.

- [Stachtari et al., 2012] Stachtari, E., Mentis, A., and Katsaros, P. (2012). Rigorous analysis of service composability by embedding ws-bpel into the bip component framework. In *IEEE International Conference on Web Services*. To appear.
- [Stahl et al., 2006] Stahl, T., Völter, M., Bettin, J., Haase, A., and Helsen, S. (2006). *Model-driven software development - technology, engineering, management*. Wiley Software Patterns Series. John Wiley.
- [Thanisch, 2000] Thanisch, P. (2000). Atomic commit in concurrent computing. *IEEE Concurrency*, 8(4):34–41.
- [Transaction Processing Performance Council, 2005] Transaction Processing Performance Council (2005). TPC benchmark C, standard specification, version 5.4.
- [Varga and Hornig, 2008] Varga, A. and Hornig, R. (2008). An overview of the OMNeT++ simulation environment. In *Proceedings of the 1st international conference on Simulation tools and techniques for communications, networks and systems & workshops, Simutools '08*, pages 1–10, ICST, Brussels, Belgium, Belgium. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).
- [Verizon, 2012] Verizon (2012). Verizon network latency statistics.
- [Vieira and Madeira, 2003] Vieira, M. and Madeira, H. (2003). A dependability benchmark for OLTP application environments. In *In proceedings of the 29th international conference on very large data bases (VLDB)*, pages 742–753.
- [Wang et al., 2008] Wang, T., Vonk, J., Kratz, B., and Grefen, P. W. P. J. (2008). A survey on the history of transaction management: from flat to grid transactions. *Distributed and Parallel Databases*, 23(3):235–270.
- [Williams and Carver, 2010] Williams, B. J. and Carver, J. C. (2010). Characterizing software architecture changes: A systematic review. *Information and Software Technology*, 52(1):31–51.

- [Wu et al., 2004] Wu, H., Kemme, B., and Maverick, V. (2004). Eager replication for stateful J2EE servers. In *In Int. Symp. on Distributed Objects and Applications (DOA)*, pages 1376–1394. Springer Verlag.
- [Yang and Dill, 1998] Yang, C. H. and Dill, D. L. (1998). Validation with guided search of the state space. In *Proceedings of the 35th annual Design Automation Conference, DAC '98*, pages 599–604, New York, NY, USA. ACM.
- [Zhu et al., 2007] Zhu, L., Bui, N. B., Liu, Y., and Gorton, I. (2007). MDABench: Customized benchmark generation using MDA. *Journal of Systems and Software*, 80(2):265–282.