

Τίτλος Έργου: «StochSocS: Συστήματα σε Ψηφίδα για Παράλληλη Στοχαστική Προσομοίωση Βιολογικών Δικτύων στη Βιολογία Συστημάτων»

(κωδικός 3828 και Κ.Α. 70/3/12367)

ΠΑΡΑΔΟΤΕΟ 1.1

Σχεδίαση Αρχιτεκτονικής SoC για τον FRM-SSA

ΑΘΗΝΑ

ΦΕΒΡΟΥΑΡΙΟΣ 2015



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην ποιότητα της γνώσης

ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΣΠΑ
2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

1.	ΕΙΣΑΓΩΓΗ.....	6
1.1	Η Σημασία της Συστημικής Βιολογίας.....	6
1.2	Η Πρόκληση.....	7
1.3	Ο Στόχος.....	8
1.4	Διάρθρωση Παραδοτέου.....	9
2.	ΜΕΘΟΔΟΙ ΣΤΟΧΑΣΤΙΚΗΣ ΠΡΟΣΟΜΟΙΩΣΗΣ.....	10
2.1	Μοντελοποίηση Βιολογικών Δικτύων.....	10
2.2	Μαρκοβιανές Διαδικασίες.....	11
2.3	Ο αλγόριθμος SSA του Gillespie.....	12
2.4	Ο αλγόριθμος FRM-SSA του Gillespie.....	14
2.5	Ο αλγόριθμος NRM-SSA των Gibson και Bruck.....	15
3.	ΠΡΟΔΙΑΓΡΑΦΕΣ ΣΥΣΤΗΜΑΤΟΣ.....	16
3.1	Προδιαγραφές αρχιτεκτονικής FRM SoC.....	16
3.2	Επικοινωνία Συστήματος.....	17
3.3	Λειτουργία Συστήματος και Ροή Εργασιών.....	18
4.	ΑΡΧΙΤΕΚΤΟΝΙΚΗ FRM ΣΥΣΤΗΜΑΤΟΣ ΣΕ ΨΗΦΙΔΑ.....	23
4.1	Επίπεδο Συστήματος - Δίκτυο FRM Πυρήνων.....	23
4.2	AXIS4 SWITCH – AXIS4 BRIDGE.....	24
4.3	Δομές Δεδομένων του FRM πυρήνα.....	26
4.4	Αρχιτεκτονική FRM Πυρήνα.....	28
4.4.1	FRM Multiprocessing Unit - FMU.....	32
4.4.2	Species/Time Update Unit - STU.....	33
4.4.3	ST Initialization Memory System – STI MS.....	35
4.4.4	Simulation Control Unit – SCU.....	36
4.4.5	Input Interface System - IIS.....	38
4.4.6	Output Interface System - OIS.....	39
4.4.7	Αναπαράσταση Αποτελεσμάτων.....	42
5.	ΑΡΧΙΤΕΚΤΟΝΙΚΗ FRM ΕΠΕΞΕΡΓΑΣΤΙΚΗΣ ΜΟΝΑΔΑΣ.....	44
5.1	FRM Processing Control Unit - FPCU.....	46
5.2	Reaction Table Memory System – RT MS.....	46
5.3	Species Table Memory System – ST MS.....	48
5.4	Processing Element - PE.....	50
5.5	Propensity Calculation Unit - PCU.....	53



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΣΠΑ
2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

5.6	Θεωρητική Ανάλυση Επιδόσεων.....	56
6.	ΣΤΑΤΙΣΤΙΚΑ ΥΛΟΠΟΙΗΣΗΣ	61
7.	ΕΠΙΚΥΡΩΣΗ ΣΧΕΔΙΑΣΗΣ	64
	ΑΝΑΦΟΡΕΣ	67



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης

ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΣΠΑ
2007-2013
Πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

ΠΙΝΑΚΑΣ ΕΙΚΟΝΩΝ

Εικόνα 1: Προγραμματισμός του FPGA με την αρχιτεκτονική FRM SoC.....	19
Εικόνα 2: Ροή εργασιών προσομοίωσης στο υπολογιστικό σύστημα Host PC και FRM SoC.....	21
Εικόνα 3: FRM-SSA SoC, αρχιτεκτονική συστήματος	23
Εικόνα 4: Μπλοκ διάγραμμα της μονάδας AXIS4 BRIDGE.....	25
Εικόνα 5: Αρχιτεκτονική του FRM πυρήνα.....	28
Εικόνα 6: Φάσεις λειτουργίας FRM πυρήνα και ροή δεδομένων κατά τη διάρκεια ενός κύκλου αντιδράσεων.	30
Εικόνα 7: Αρχιτεκτονική της μονάδας FRM Multiprocessing Unit - FMU.....	32
Εικόνα 8: Διάγραμμα κυματομορφών της μονάδας SPU.....	34
Εικόνα 9: Αρχιτεκτονική της μονάδας STU.....	35
Εικόνα 10: Μπλοκ διάγραμμα της μονάδας STI MS	36
Εικόνα 11: Μπλοκ διάγραμμα SCU μονάδας	37
Εικόνα 12: Μπλοκ διάγραμμα της μονάδας IIS	38
Εικόνα 13: Αρχιτεκτονική της μονάδας OIS.....	39
Εικόνα 14: Διάγραμμα μηχανισμού δειγματοληψίας.....	41
Εικόνα 15: Αρχιτεκτονική της μονάδας FPU.....	44
Εικόνα 16: Μπλοκ διάγραμμα της μονάδας FPCU	46
Εικόνα 17: Αρχιτεκτονική της μονάδας RT MS	47
Εικόνα 18: Αρχιτεκτονική της μονάδας ST MS.....	49
Εικόνα 19: Αρχιτεκτονική της μονάδας PE	51
Εικόνα 20: Αρχιτεκτονική της μονάδας Propensity Calculation Unit - PCU.....	54
Εικόνα 21: Θεωρητικά αναμενόμενη διεκπεραιωτική ικανότητα και επίδοση του συστήματος FRM SoC ως συνάρτηση του πλήθους αντιδράσεων του βιομοντέλου (m) και του αριθμού των PEs (N). Συχνότητα λειτουργίας 175 MHz.....	59
Εικόνα 22: Αποτελέσματα προσομοιώσεων με το βιομοντέλο LVS	65
Εικόνα 23: Σύγκριση των προσομοιωτών FRM NoC και Copasi	65



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης

ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΣΠΑ
2007-2013
Πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

ΠΙΝΑΚΑΣ ΠΙΝΑΚΩΝ

Πίνακας 1: Περιγραφή των σημάτων του πρωτοκόλλου AXIS4	25
Πίνακας 2: Η δομή των πινάκων RT και ST του FRM πυρήνα	27
Πίνακας 3: Η ακολουθία δεδομένων που παράγεται από την μονάδα STU.....	31
Πίνακας 4: Τύποι και Ετικέτες Δεδομένων	42
Πίνακας 5: Υπολογισμός των τάσεων (propensities) των αντιδράσεων έως τρίτης τάξης	54
Πίνακας 6: Λειτουργίες της υπομονάδας Reaction Order SWITCH της PCU.....	55
Πίνακας 7: Καθυστερήσεις λόγω διοχέτευσης της PE μονάδας	57
Πίνακας 8: Καθυστερήσεις λόγω διοχέτευσης σε επίπεδο FRM πυρήνα	57
Πίνακας 9: Καθυστερήσεις λόγω διοχέτευσης της διόδου δεδομένων	58
Πίνακας 10: Χαρακτηριστικά του FPGA XC7K325T	61
Πίνακας 11: Στατιστικά υλοποίησης του FRM πυρήνα.....	61
Πίνακας 12: Στατιστικά υλοποίησης του PCIe πυρήνα και της μονάδας AXIS4 SWITCH.....	62
Πίνακας 13: Στατιστικά υλοποίησης του FRM1xN SoC συστήματος - SSIP	62
Πίνακας 14: Στατιστικά υλοποίησης του FRMCx1 SoC συστήματος - MSIP	62



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης

ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΣΠΑ
2007-2013
Πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

1. ΕΙΣΑΓΩΓΗ

1.1 Η Σημασία της Συστημικής Βιολογίας

Τη τελευταία δεκαετία έχει γίνει σαφές ότι η *Συστημική Βιολογία* (Systems Biology) διαδραματίζει καθοριστικό ρόλο στην μελέτη και στην ερμηνεία διάφορων πολύ σημαντικών βιολογικών συστημάτων [1]. Η συστημική προσέγγιση έχει προφανή πλεονεκτήματα στη μελέτη της βιολογίας καθώς είναι πλέον αποδεκτό ότι δεν αρκεί να μελετάμε απομονωμένα τμήματα ενός συστήματος εφόσον αυτά αλληλεπιδρούν και αλληλοεπηρεάζονται [2]. Μπορούμε να διακρίνουμε το κύτταρο καθώς και άλλα επίπεδα οργάνωσης (ιστούς, όργανα) ως ένα συνολικό σύστημα του οποίου τα μέρη δε λειτουργούν αυτόνομα καθώς αλληλοεξαρτώνται και ρυθμίζουν τις λειτουργίες τους αποκρινόμενα στις απαιτήσεις και στις αλλαγές του περιβάλλοντος, όλα μαζί σε συνεργασία. Ένα ακόμη πλεονέκτημα, αυτής της ολιστικής προσέγγισης, είναι ότι η έννοια του συστήματος χρησιμοποιείται εδώ και πάρα πολλά χρόνια από πολλές διαφορετικές επιστήμες με ισχυρό μαθηματικό υπόβαθρο, έτσι λοιπόν, η εισαγωγή μιας τέτοιας προσέγγισης στη βιολογία δίνει νέα ώθηση στη βαθύτερη κατανόηση της.

Ωστόσο, η μελέτη ενός βιολογικού συστήματος οδηγεί στη συσσώρευση τρομακτικού όγκου δεδομένων και πληροφορία. Κάτι τέτοιο καθιστά αναγκαία και απαραίτητη τη χρήση υπολογιστικών μεθόδων. Στην αντιμετώπιση του τεράστιου όγκου δεδομένων συνδράμουν οι κλάδοι της Υπολογιστικής Βιολογίας (Computational Biology) [3] και της Βιοπληροφορική (Bioinformatics) [4]. Η Υπολογιστική Βιολογία ασχολείται με την ανάπτυξη θεωρητικών μεθόδων και μαθηματικών μοντέλων για τη μελέτη συστημάτων βιολογίας σε διάφορα επίπεδα πολυπλοκότητας. Ενώ η Βιοπληροφορική ασχολείται με την έρευνα και ανάπτυξη κατάλληλων υπολογιστικών εργαλείων και προσεγγίσεων που στοχεύουν σε ένα ευρύ φάσμα εφαρμογών με σκοπό την αποδοτικότερη μελέτη και ανάλυση των παραγόμενων δεδομένων. Για το σκοπό αυτό, η Υπολογιστική Βιολογία και η Βιοπληροφορική χρησιμοποιούν μεθόδους κλάδων όπως η ανάλυση αλγορίθμων, οι βάσεις δεδομένων, η εξόρυξη δεδομένων [5], η υπολογιστική θεωρία, η θεωρία ελέγχου [6], η θεωρία συστημάτων, η θεωρία κυκλωμάτων [7], η μοντελοποίηση και προσομοίωση συστημάτων, η τεχνητή νοημοσύνη [8], η αναγνώριση προτύπων [9], η επεξεργασία σημάτων, η ανάπτυξη λογισμικού, η ανάπτυξη συστημάτων κτλ.

Η Βιοπληροφορική θεωρώντας τα βιολογικά δεδομένα ως ψηφιακή πληροφορία έχει καταφέρει να παίζει καθοριστικό ρόλο στη μελέτη της βιολογίας. Είναι ευρέως αποδεκτό ότι χωρίς τη συνδρομή των υπολογιστικών μεθόδων θα ήταν αδύνατο να αναλυθεί ο τεράστιος όγκος δεδομένων που παράγει η μελέτη ενός βιολογικού συστήματος και τελικά να εξαχθούν χρήσιμες πληροφορίες για τη δυναμική λειτουργία του [10].

Χάρη στη συστημική προσέγγιση μπορούμε να μοντελοποιήσουμε με δίκτυα διάδοσης της πληροφορίας το κύτταρο αλλά και επίπεδα οργάνωσης του [11]. Η προσέγγιση αυτή διαφοροποιεί αρκετά τη συστημική βιολογία από τη παραδοσιακή μέθοδο. Ένα πλεονέκτημα της προσέγγισης αυτής είναι ότι δε γίνεται απλά μια συσχέτιση των βιολογικών συμβάντων αλλά δομείται ένα ολόκληρο σύστημα όπου σε αυτό υπεισέρχεται και η επικοινωνία μεταξύ των μερών καθώς και τα αποτελέσματα που έχει εν τέλει η ενεργοποίηση κάθε μέρους ξεχωριστά.

Ένα ακόμη πλεονέκτημα της προσέγγισης αυτής είναι ότι πλέον είναι εφικτή υπολογιστικά η ανακατασκευή αλλά και η προσομοίωση μεγάλων δικτύων αλληλεπιδράσεων που μοντελοποιούν τη συμπεριφορά ολόκληρων κυτταρικών υποσυστημάτων. Η δυνατότητα προσομοίωσης με υπολογιστές βιολογικών συστημάτων έχει αποδειχθεί ως ένα εξαιρετικά χρήσιμο και απαραίτητο εργαλείο για τη μελέτη της δυναμικής συμπεριφοράς των κυτταρικών μηχανισμών [12]. Εδώ αξίζει να σημειωθεί ότι τα τελευταία χρόνια έχει αναδειχθεί σε ζωτικής σημασίας η μοντελοποίηση του κυττάρου σε ένα σύστημα [13]. Υπάρχει μεγάλη ελπίδα ότι αν καταφέρουμε να μοντελοποιήσουμε το κύτταρο σαν ένα σύστημα τότε θα βρεθούν θεραπείες για πολλές ανίατες μέχρι τώρα ασθένειες.

Όλες οι προαναφερθείσες προσεγγίσεις ανοίγουν νέους ορίζοντες στη μελέτη της βιολογίας. Ο συνδυασμός της συστημικής και υπολογιστικής βιολογίας δίνουν τη δυνατότητα εκτέλεσης χιλιάδων πειραμάτων σε υπολογιστές με λογικό κόστος, απαντώντας σε πολύ κρίσιμα ερωτήματα σε πολύ μικρό χρονικό διάστημα. Ένα ακόμη χαρακτηριστικό της τάσης αυτής είναι ότι πλέον τα πειράματα δεν καθορίζονται από τους ίδιους τους βιολόγους αλλά από υπολογιστικά συστήματα που έχουν κάνει όχι μόνο την ανάλυση του τεράστιου όγκου δεδομένων αλλά έχουν καταλήξει και στα επόμενα πειράματα που έχει νόημα να διεξαχθούν. Κάτι τέτοιο, όπως είναι σαφές, δίνει τη δυνατότητα στους βιολόγους να επικεντρωθούν μόνο στη παρακολούθηση των βιολογικών πειραμάτων και στη κατανόηση τους. Το αποτέλεσμα όλων των παραπάνω είναι η ταχεία συσσώρευση γνώσης αλλά και η δημιουργία καινούργιων αναπάντητων ερωτημάτων που αξίζει να διερευνηθούν.

1.2 Η Πρόκληση

Η κύρια πρόκληση στη προσπάθεια μοντελοποίησης του κυττάρου με τη χρήση βιολογικών δικτύων είναι η αντιμετώπιση της πολυπλοκότητας του μοντέλου και της στοχαστικότητας του μηχανισμού που περιγράφει. Όσον αφορά τη πολυπλοκότητα του μοντέλου, χαρακτηριστικό παράδειγμα είναι το μεταβολικό δίκτυο E-coli [14]. Στο παράδειγμα αυτό αρκεί να αναφερθεί ότι ο αριθμός των βιοχημικών αντιδράσεων που πραγματοποιούνται στη διάρκεια ενός κυτταρικού κύκλου υπολογίζεται ότι είναι της τάξης 10^{14} έως 10^{16} . Οι πλατφόρμες λογισμικού που έχουν αναπτυχθεί για τη προσομοίωση βιολογικών συστημάτων δε πετυχαίνουν υψηλότερη απόδοση από 0.5 Mreactions/sec (εκατομμύρια αντιδράσεις το δευτερόλεπτο) ακόμη και αν χρησιμοποιηθούν τελευταίας γενιάς επεξεργαστικές μονάδες (CPUs/GPUs) για την εκτέλεση των προγραμμάτων προσομοίωσης [15] [16]. Από τη παραπάνω σύγκριση απαιτήσεων και διαθέσιμων υπολογιστικών πόρων, καθίσταται σαφές ότι η προσομοίωση ολόκληρων κυττάρων θα παραμείνει άπιαστος στόχος εκτός και εάν αναπτυχθούν νέα υπολογιστικά συστήματα και εξειδικευμένο υλικό τα οποία συνδυαζόμενα με νέους πιο αποδοτικούς αλγόριθμους καταφέρουν να καλύψουν τις απαιτήσεις της σύγχρονης βιολογίας.

Η παραδοσιακή μέθοδος που συνεχίζει να χρησιμοποιείται για την περιγραφή της δυναμικής συμπεριφοράς βιολογικών συστημάτων στηρίζεται στη χρήση Συνήθων Διαφορικών Εξισώσεων (Ordinary Differential Equations - ODEs). Ωστόσο σε ένα μοντέλο που περιγράφεται από ODEs έχουν γίνει δύο πολύ βασικές παραδοχές που δε συναντάμε στη φύση [17]. Η πρώτη παραδοχή είναι η ντετερμινιστική συμπεριφορά ενός συστήματος. Η δεύτερη παραδοχή σχετίζεται με τη ποσοτικοποίηση των μοριακών ειδών. Ωστόσο, ένα πραγματικό βιοχημικό σύστημα δεν ικανοποιεί καμία από τις παραπάνω υποθέσεις. Από τις σύγχρονες θεωρίες κβαντομηχανικής γνωρίζουμε ότι σε καμία περίπτωση ένα πραγματικό βιοχημικό σύστημα δε μπορεί να θεωρηθεί ντετερμινιστικό. Όσον αφορά τη δεύτερη παραδοχή, στη φύση, ο αριθμός των μοριακών ειδών είναι

μια διακριτή ποσότητα και πάντοτε είναι ένας φυσικός αριθμός. Επιπλέον, είναι γεγονός ότι η δυναμική συμπεριφορά ενός συστήματος μπορεί να επηρεαστεί σημαντικά από σχετικά μικρούς μοριακούς πληθυσμούς. Είναι λοιπόν πλέον ευρέως αποδεκτό να χρησιμοποιείται και στοχαστική προσέγγιση για τη μελέτη και τη μοντελοποίηση ενός βιολογικού συστήματος. Ένα ακόμη πλεονέκτημα των στοχαστικών διαδικασιών, σε σύγκριση με τη χρήση των ODEs, είναι ότι μπορούν να οδηγήσουν στην επινόηση αλγορίθμων με αποτέλεσμα τη πιο αποδοτική χρήση των υπολογιστικών συστημάτων για την προσομοίωση βιολογικών συστημάτων.

Για όλους τους παραπάνω λόγους, Μαρκοβιανές μέθοδοι, όπως οι αλγόριθμοι στοχαστικής προσομοίωσης του Gillespie (FRM-SSA) [17], έχουν αναδειχθεί ως η μέθοδος επιλογής για τη μελέτη και τη μοντελοποίηση ενός βιολογικού συστήματος. Οι αλγόριθμοι αυτοί, πέρα του ότι είναι υπολογιστικά αποδοτικοί, έχουν αποδειχθεί ότι μπορούν να περιγράψουν και να προσεγγίσουν με μεγάλη ακρίβεια τη συμπεριφορά ενός βιολογικού δικτύου.

Συνεπώς, η πρόκληση στη ανάπτυξη συστημάτων για τη στοχαστική προσομοίωση βιολογικών δικτύων εκτείνεται σε δύο διαστάσεις. Πρώτον, θα πρέπει να είναι σε θέση να προσομοιώνουν αποδοτικά πολύπλοκα βιολογικά δίκτυα χιλιάδων αντιδράσεων και μοριακών ειδών. Η δεύτερη διάσταση της πρόκλησης σχετίζεται με το γεγονός ότι η στοχαστική προσομοίωση ενός συστήματος προϋποθέτει την διεξαγωγή χιλιάδων επαναλήψεων του ίδιου πειράματος.

1.3 Ο Στόχος

Αναλογιζόμενοι την ανάγκη για αποδοτικά υπολογιστικά συστήματα που θα είναι σε θέση να προσομοιώνουν εκατομμύρια αντιδράσεις το δευτερόλεπτο ξεκινήσαμε την έρευνα μας. Μελετώντας εκτενώς τη σύγχρονη βιβλιογραφία για τις ήδη υπάρχουσες FRM αρχιτεκτονικές και με γνώμονα την αποδοτική διεξαγωγή in silico πειραμάτων μέσω στοχαστικών προσομοιώσεων σχεδιάσαμε, αναπτύξαμε και επικυρώσαμε μια πολυπύρηνη, παράλληλη αρχιτεκτονική (FRM SoC) για τη παραλληλοποίηση του αλγορίθμου Gillespie FRM-SSA [18] [19] [20] [21] [22] [23] [24] [25] [26] [27]. Ο διπλός στόχος της παρούσας έρευνας ήταν η ανάπτυξη ενός συστήματος το οποίο θα αντιμετωπίζει όλες τις προαναφερθείσες προκλήσεις και θα αυξάνει τουλάχιστον κατά μια τάξη μεγέθους τη πολυπλοκότητα των βιολογικών συστημάτων που μπορούν να προσομοιωθούν στοχαστικά. Ενώ παράλληλα, το σύστημα που θα υλοποιηθεί να είναι πλήρως παραμετροποιήσιμο ανάλογα με τα χαρακτηριστικά ενός βιολογικού δικτύου και των διαθέσιμων πόρων υλικού. Η ευέλικτη FRM SoC αρχιτεκτονική είναι παραμετρική ως προς τον αριθμό των μοριακών ειδών και τον αριθμό των αντιδράσεων ενώ καθορίζεται από τον χρήστη, κατά το στάδιο της σύνθεσης, ο αριθμός των πυρήνων και των επεξεργαστικών μονάδων που θα απαρτίζουν το τελικό σύστημα.

Ένας επιπλέον στόχος που τέθηκε κατά το στάδιο της σχεδίασης ήταν το σύστημα που θα αναπτυχθεί να είναι σε θέση να προσομοιώσει παράλληλα διαφορετικά βιολογικά δίκτυα, τελείως ανεξάρτητα μεταξύ τους, κάτι το οποίο επιτυγχάνεται μέσω της πολυπύρηνης αρχιτεκτονικής. Κάθε πυρήνας της αρχιτεκτονικής λειτουργεί ως ένα πλήρως ανεξάρτητο και παραλληλοποιημένο σύστημα προσομοίωσης του αλγορίθμου FRM-SSA.

Ευελπιστούμε ότι η έρευνα αυτή θα αποτελέσει ένα αποφασιστικό βήμα προς τη κατεύθυνση της προσομοίωσης ολόκληρου του κυττάρου, μια από τις μεγαλύτερες προκλήσεις του 21^{ου} αιώνα στο τομέα της υπολογιστικής βιολογίας [13].

1.4 Διάρθρωση Παραδοτέου

Το υπόλοιπο της αναφοράς οργανώνεται ως εξής: Στην Ενότητα 2 παρουσιάζονται οι στοχαστικές μέθοδοι προσομοίωσης και αλγόριθμοι SSA και FRM-SSA του Gillespie.

Στην Ενότητα 3 παρουσιάζονται αναλυτικά οι προδιαγραφές του συστήματος που υλοποιήθηκε, ο βαθμός παραμετροποίησης του καθώς και οι τρόποι λειτουργίας του.

Στην Ενότητα 4 αναλύεται η αρχιτεκτονική FRM SoC σε επίπεδο συστήματος καθώς επίσης γίνεται και σύντομη αναφορά στο σύστημα επικοινωνίας υπολογιστή και συστήματος.

Στην Ενότητα 5 παρουσιάζεται η αρχιτεκτονική της επεξεργαστικής μονάδας (FRM Processing Unit - FPU) ενός SSA Core. Δίνεται έμφαση στη δίοδο δεδομένων της FPU ενώ περιγράφονται αναλυτικά και οι υπόλοιπες μονάδες που πλαισιώνουν τη δίοδο δεδομένων της FPU. Επιπλέον παρουσιάζεται και η θεωρητική μελέτη των επιδόσεων που έγινε κατά το σχεδιασμό.

Στην Ενότητα 6 παρουσιάζονται τα στατιστικά αποτελέσματα που προέκυψαν από τη σύνθεση του συστήματος για διάφορους τρόπους λειτουργίας.

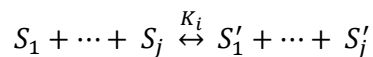
Στην 7 και τελευταία ενότητα παρουσιάζονται πραγματικά αποτελέσματα από δοκιμές του συστήματος με σκοπό την επικύρωση της σχεδίασης. Για αυτό το λόγο γίνεται σύγκριση των αποτελεσμάτων με τα αποτελέσματα γνωστών πλατφόρμων προσομοίωσης.

2. ΜΕΘΟΔΟΙ ΣΤΟΧΑΣΤΙΚΗΣ ΠΡΟΣΟΜΟΙΩΣΗΣ

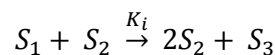
Σε αυτή την ενότητα αρχικά παρουσιάζεται το θεωρητικό πλαίσιο που απαιτείται για τη μοντελοποίηση ενός βιολογικού συστήματος. Έπειτα ακολουθεί η στοχαστική ανάλυση ενός συστήματος μέσω Μαρκοβιανών αλυσίδων και τα πλεονεκτήματα της εν αντιθέσει με τη ντετερμινιστική προσέγγιση των ODEs. Στις δύο τελευταίες ενότητες παρουσιάζονται εκτενώς οι στοχαστικές μέθοδοι προσομοίωσης Stochastic Simulation Algorithm-SSA και First Reaction Method-FRM του Gillespie και γίνεται αναφορά στη μέθοδο στοχαστικής προσομοίωσης Next Reaction Method-NRM των Gibson και Bruck.

2.1 Μοντελοποίηση Βιολογικών Δικτύων

Στη βιολογία συστημάτων η μοντελοποίηση ενός βιολογικού δικτύου γίνεται μέσω βιομοντέλων. Τα βιομοντέλα αποτελούνται που μοντελοποιούν τη συμπεριφορά του βιολογικού συστήματος. Στη συνέχεια ακολουθεί η τυπική μορφή μιας αντίδρασης R_i ενός βιομοντέλου:



Με S_j συμβολίζονται τα μοριακά είδη που συμμετέχουν σε μια αντίδραση. Σε στοχαστικές προσομοιώσεις τα μοριακά είδη είναι μια διακριτή ποσότητα και συμβολίζουν μοριακούς πληθυσμούς του είδους j . Τα αριστερό μέρος της αντίδρασης καθορίζει τα μοριακά είδη S_j που συμμετέχουν ως αντιδρώντα σε αυτή και αντίστοιχα το δεξιό μέρος της καθορίζει τα μοριακά είδη S'_j που θα παραχθούν από την αντίδραση, δηλαδή τα προϊόντα. Με τη σταθερά K_i μοντελοποιείται ο ρυθμός της αντίδρασης ενώ η κατεύθυνση του βέλους καθορίζει την αντιστρεψιμότητα της αντίδρασης. Τέλος, ο αριθμός των μοριακών ειδών που συμμετέχουν ως αντιδρώντα σε μια αντίδραση καθορίζει τη τάξη της αντίδρασης. Στη συνέχεια δίνεται ένα απλό παράδειγμα μιας αντίδρασης που διέπεται από τους νόμους της κινητικής δευτέρας τάξης:

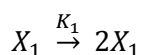


Η παραπάνω σχέση διαβάζεται ως εξής: ένα μοριακό είδος S_1 θα αντιδράσει με ένα μοριακό είδος S_2 , με ρυθμό K_i , και από την αντίδραση τους θα παραχθούν δύο μοριακά είδη S_2 και ένα μοριακό είδος S_3 . Ο λόγος που τελικά παράγεται ένα μοριακό είδος S_2 είναι ότι το συγκεκριμένο είδος συμμετέχει και ως αντιδρόν και ως προϊόν στην αντίδραση. Ο συντελεστής 2 μπροστά από το είδος S_2 αλλά και ο συντελεστής 1, που παραλείπεται μπροστά από τα υπόλοιπα είδη, καθορίζουν τη ποσότητα των μοριακών ειδών που συμμετέχουν στην αντίδραση. Η ποσότητα των μοριακών ειδών που συμμετέχουν σε μια αντίδραση ονομάζεται *στοιχειομετρία* της

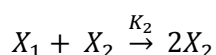
αντίδρασης. Συνεπώς το διάνυσμα της στοιχειομετρίας της παραπάνω αντίδρασης είναι $[1 \ 1 \ 2 \ 1]$. Η δεξιά κατεύθυνση του βέλους δείχνει ότι η αντίδραση είναι μη αντιστρέψιμη.

Για τη καλύτερη κατανόηση το πώς μοντελοποιείται ένα βιολογικό σύστημα μέσω αντιδράσεων δίνεται ένα απλό παράδειγμα συστήματος τύπου θηρευτή/θηραμάτος. Το σύνολο αντιδράσεων που ακολουθεί μοντελοποιεί το σύστημα Lotka Voltera [28] το οποίο περιγράφεται από τρεις αντιδράσεις και δύο μοριακά είδη. Με X_1 συμβολίζεται ο μοριακός πληθυσμός των θηραμάτων και με X_2 ο μοριακός πληθυσμός των θηρευτών.

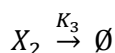
- R_1 , αύξηση του πληθυσμού των θηραμάτων



- R_2 , αύξηση του πληθυσμού των θηρευτών και μείωση του πληθυσμού των θηραμάτων



- R_3 , μείωση του πληθυσμού των θηρευτών



Η ανάγκη για ανταλλαγή βιομοντέλων μεταξύ της επιστημονικής κοινότητας δημιούργησε την απαίτηση για έναν ενιαίο τρόπο κωδικοποίησης βιομοντέλων που θα συμμορφώνονται με τις προδιαγραφές ενός κοινού προτύπου. Για το λόγο αυτό αναπτύχθηκε μια γλώσσα περιγραφής βιομοντέλων η οποία ονομάστηκε SBML (System Biology Markup Language) [29] και στηρίχτηκε στη γλώσσα σήμανσης και κωδικοποίησης ηλεκτρονικών κειμένων XML (Extensible Markup Language) [30]. Μέσω της SBML μπορεί να κωδικοποιηθεί ένα ολόκληρο βιολογικό σύστημα και στη συνέχεια να διατίθεται μέσω ηλεκτρονικών βάσεων δεδομένων σε όλη την επιστημονική κοινότητα. Η SBML δίνει τη δυνατότητα κωδικοποίησης των αντιδράσεων, των μοριακών ειδών, των σταθερών και των παραμέτρων καθώς και των μαθηματικών μοντέλων και κανόνων που περιγράφουν ένα βιολογικό σύστημα σε αυτό που ονομάζεται SBML μοντέλο.

2.2 Μαρκοβιανές Διαδικασίες

Όπως αναφέρθηκε και στην εισαγωγική ενότητα, η παραδοσιακή μέθοδος που χρησιμοποιείται για τη περιγραφή της δυναμικής συμπεριφοράς ενός βιολογικού συστήματος στηρίζεται σε ντετερμινιστικές μεθόδους οι οποίες περιγράφονται με τη χρήση Συνήθων Διαφορικών Εξισώσεων (ODEs). Αυτές οι μέθοδοι καθορίζουν αυστηρώς την μελλοντική κατάσταση του συστήματος κάθε χρονική στιγμή t , με κατάλληλες αρχικές και συνοριακές συνθήκες για οποιαδήποτε κατάσταση του συστήματος στο παρόν, παρελθόν ή μέλλον [31]. Ωστόσο, από τη κβαντομηχανική αποδεικνύεται ότι οι ντετερμινιστικές μέθοδοι δε μπορούν να περιγράψουν ικανοποιητικά ένα πραγματικό σύστημα λόγω της στοχαστικότητας που απαντάται στη φύση ενώ παράλληλα

δεν είναι υπολογιστικά αποδοτικές [32]. Για τους προαναφερθέντες λόγους οι στοχαστικές μέθοδοι προτιμούνται για τη περιγραφή και τη μελέτη της συμπεριφοράς δυναμικών συστημάτων.

Αντίθετα με τις ντετερμινιστικές μεθόδους, στις στοχαστικές διαδικασίες μπορεί να υπολογιστεί μόνο η πιθανότητα του να βρεθεί το σύστημα σε μια συγκεκριμένη κατάσταση σε συγκεκριμένο χρόνο. Επιπλέον, εάν η πιθανότητα μετάβασης του συστήματος σε μια συγκεκριμένη κατάσταση εξαρτάται μόνο από τη κατάσταση που ήδη βρίσκεται και όχι από τις προηγούμενες καταστάσεις τότε το σύστημα μπορεί να μοντελοποιηθεί με Μαρκοβιανές διαδικασίες [33].

Μέσω της πιθανοθεωρίας και των Μαρκοβιανών αλυσίδων μπορούμε να μελετήσουμε την εξέλιξη ενός δυναμικού συστήματος με τη χρήση θεμελιωδών γεγονότων μετάβασης από τη μια κατάσταση σε κάποια άλλη. Κάθε θεμελιώδες γεγονός περιγράφεται ως η μετάβαση από τη παρούσα κατάσταση w του συστήματος προς μια καινούρια κατάσταση w' . Έστω ότι $P(w, t)$ η πιθανότητα να βρίσκεται το σύστημα στη κατάσταση w τη χρονική στιγμή t , αποδεικνύεται ότι η εξέλιξη της πιθανότητας $P(w, t)$ διέπεται από την ακόλουθη εξίσωση:

(1)

$$\frac{\partial P(w, t)}{\partial t} = \sum_{w' \in W, w' \neq w} T(w' \rightarrow w)P(w', t) - T(w \rightarrow w')P(w, t)$$

όπου το άθροισμα εκτείνεται σε όλο το χώρο καταστάσεων του συστήματος και είναι ικανό να περιγράψει τη δυναμική εξέλιξη ενός στοχαστικού συστήματος κάθε κατάστασης w . Η συνάρτηση μετάβασης $T(w \rightarrow w')$ δίνει τη πυκνότητα πιθανότητας ανά μονάδα χρόνου να μεταβεί το σύστημα από τη κατάσταση w στη κατάσταση w' . Η παραπάνω εξίσωση ονομάζεται Chemical Master Equation – CME [17], ο πρώτος όρος της εξίσωσης ουσιαστικά εκφράζει τη πιθανότητα το σύστημα να μεταβεί από άλλες καταστάσεις w' στη παρούσα κατάσταση w ενώ ο δεύτερος όρος της εξίσωσης τη πιθανότητα το σύστημα να μεταβεί από τη παρούσα κατάσταση w σε άλλες καταστάσεις w' . Η επίλυση της CME μας δίνει τη πιθανότητα το σύστημα να βρίσκεται στη κατάσταση w τη χρονική στιγμή t . Δυστυχώς όμως, η CME δε δίνει αναλυτικές ή αριθμητικές λύσεις παρά μόνο σε ελάχιστες περιπτώσεις πολύ απλών συστημάτων, αφού θα πρέπει να επιλυθεί για κάθε μοριακό είδος ενός βιομοντέλου. Επιπλέον η προσπάθεια επίλυσης ενός σετ από CMEs είναι υπολογιστικά μη αποδοτική έως ανέφικτη και για αυτό κυρίως χρησιμοποιείται κατά το στάδιο αξιολόγησης στοχαστικών αλγορίθμων [34].

2.3 Ο αλγόριθμος SSA του Gillespie

Παρόλο που συνήθως η CME δεν μπορεί να δώσει αναλυτικές ή αριθμητικές λύσεις ένα δυναμικό σύστημα μπορεί να μελετηθεί χρησιμοποιώντας Monte Carlo τεχνικές στοχαστικής προσομοίωσης [35]. Το μεγάλο πλεονέκτημα των Monte Carlo τεχνικών σε σύγκριση με τις προαναφερθείσες προσεγγίσεις είναι ότι εκμεταλλεύονται το γεγονός ότι η δυναμική εξέλιξη του συστήματος, σε βάθος χρόνου, καθορίζεται πρακτικά από μεμονωμένες μεταπηδήσεις από κατάσταση σε κατάσταση. Για αυτό το λόγο, οι τεχνικές Monte Carlo επικεντρώνονται μόνο σε αυτά τα σπάνια γεγονότα αγνοώντας την ενδιάμεση "περιπλάνηση" του συστήματος δίνοντας έτσι τη δυνατότητα προσομοίωσης αντιπροσωπευτικών χρονικών στιγμών.

Συνεπώς, σε ένα βιολογικό σύστημα, με τεχνικές Monte Carlo δε χρειάζεται να παρακολουθούμε το σύστημα σε κάθε αυθαίρετο χρονικά βήμα του αλλά μόνο στις χρονικές στιγμές όπου οι αντιδράσεις του μοντέλου

ενεργοποιούνται και πρακτικά μεταβάλλουν τη κατάσταση του συστήματος. Σε αντιστοιχία με τις Μαρκοβιανές αλυσίδες, πλέον ως θεμελιώδες γεγονός μπορεί να οριστεί μόνο η αυξομείωση των πληθυσμών των μοριακών ειδών που περιγράφεται από τη στοιχειομετρία μιας αντίδρασης. Ενώ ως καταστάσεις μετάβασης του συστήματος ορίζονται οι πληθυσμοί των μοριακών ειδών πριν και μετά την αυξομείωση τους.

Χρησιμοποιώντας τις παραπάνω παρατηρήσεις ο Gillespie επινόησε τον αλγόριθμο Stochastic Simulation Algorithm – SSA, ο οποίος ήταν ο πρώτος αλγόριθμος που χρησιμοποιήθηκε για τη στοχαστική προσομοίωση βιολογικών συστημάτων [36]. Στη συνέχεια ακολουθούν τα βήματα εκτέλεσης του αλγορίθμου:

1. Αρχικοποίηση του συστήματος $x = x_0$ (καθορισμός των αρχικών πληθυσμών κάθε μοριακού είδους) και αρχικοποίηση του χρόνου προσομοίωσης $t = t_0$.
2. Υπολογισμός όλων των τάσεων (propensities) $\alpha_j(x)$ των αντιδράσεων $\{R_j, j = 1, 2, \dots, m\}$ του μοντέλου στη κατάσταση x και στη συνέχεια υπολογισμός του αθροίσματος των τάσεων:

$$\alpha_0(x) = \sum_{j=1}^m \alpha_j(x)$$

3. Επιλογή δύο ψευδοτυχαίων αριθμών r_1 και r_2 στο διάστημα $[0, 1]$.
4. Υπολογισμός του χρόνου τ που θα μεσολαβήσει μέχρι την ενεργοποίηση της επόμενης αντίδρασης:

$$\tau = \frac{1}{\alpha_0(x)} \ln\left(\frac{1}{r_1}\right)$$

5. Εύρεση της αντίδρασης R_μ με το μικρότερο χρόνο:

$$\mu = \text{o μικρότερος αριθμός που ικανοποιεί τη συνθήκη } \sum_{j=1}^{\mu} \alpha_j(x) > r_2 \alpha_0(x)$$

6. Εκτέλεση της αντίδρασης R_μ και ανανέωση της κατάστασης του συστήματος $x = x + \nu_\mu$ (δηλαδή μεταβολή των πληθυσμών των μοριακών ειδών βάση του διανύσματος στοιχειομετρίας ν_μ της αντίδρασης R_μ) καθώς και του χρόνου προσομοίωσης, $t = t + \tau$.
7. Εάν το σύστημα έχει φτάσει τον επιθυμητό χρόνο προσομοίωσης (δηλαδή $t > T_{sim}$) τότε ο αλγόριθμος τερματίζει αλλιώς μεταβαίνει ξανά στο βήμα 2.

Χάρη στον αλγόριθμο SSA πλέον η μελέτη της συμπεριφοράς ενός βιολογικού συστήματος δεν απαιτεί πολύπλοκους υπολογισμούς σε αντίθεση με τη χρήση των ODEs ή τη χρήση της CME. Ωστόσο η προσέγγιση αυτή έχει μεγάλο υπολογιστικό κόστος για δύο λόγους. Ο πρώτος λόγος είναι ότι η προσομοίωση θα πρέπει να εκτελεστεί πάρα πολλές φορές, τουλάχιστον χίλιες, προκειμένου να προσεγγιστεί στατιστικά με ακρίβεια η συμπεριφορά του συστήματος [34]. Ο δεύτερος λόγος είναι ότι τα χρονικά βήματα του συστήματος μπορούν να προκύψουν πολύ μικρά εάν οι πληθυσμοί των μοριακών ειδών αυξηθούν αρκετά κάτι το οποίο συνεπάγεται επιπλέον υπολογιστική προσπάθεια.

2.4 Ο αλγόριθμος FRM-SSA του Gillespie

Ο αλγόριθμος SSA αναφέρεται και ως η "άμεση μέθοδος", Direct Method – SSA. Ο λόγος είναι ότι κάθε φορά υπολογίζεται το χρονικό βήμα τ ολόκληρου του συστήματος. Ο Gillespie πρότεινε και έναν δεύτερο αλγόριθμο όπου υπολογίζονται οι χρόνοι όλων των αντιδράσεων και κατόπιν η αντίδραση με το μικρότερο χρόνο είναι αυτή που επιλέγεται να εκτελεστεί (νικητήρια αντίδραση) [17]. Ο αλγόριθμος αυτός ονομάζεται First Reaction Method (FRM-SSA). Στη συνέχεια ακολουθούν τα βήματα εκτέλεσης του αλγορίθμου:

1. Αρχικοποίηση του συστήματος $x = x_0$ (καθορισμός των αρχικών πληθυσμών κάθε μοριακού είδους) και αρχικοποίηση του χρόνου προσομοίωσης, $t = t_0$.
2. Υπολογισμός όλων των τάσεων (propensities) $\alpha_j(x)$ των αντιδράσεων $\{R_j, j = 1, 2, \dots, m\}$ του μοντέλου.
3. Για κάθε αντίδραση R_j , υπολογισμός του χρόνου της δυνητικής εκτέλεσης:

$$\tau_j = \frac{1}{\alpha_j(x)} \ln\left(\frac{1}{r_j}\right)$$

όπου r_j είναι ψευδοτυχαίοι μοναδιαίοι αριθμοί στο διάστημα $[0, 1]$

4. Εύρεση της αντίδρασης R_μ με το μικρότερο χρόνο ενεργοποίησης τ_μ .
5. Εκτέλεση της αντίδρασης R_μ ανανεώνοντας τη κατάσταση του συστήματος $x = x + \nu_\mu$ (δηλαδή μεταβολή των πληθυσμών των μοριακών ειδών βάση του διανύσματος στοιχειομετρίας ν_μ της αντίδρασης R_μ) καθώς και του χρόνου προσομοίωσης, $t = t + \tau_\mu$.
6. Εάν το σύστημα έχει φτάσει τον επιθυμητό χρόνο προσομοίωσης ($t > T_{sim}$) τότε ο αλγόριθμος τερματίζει αλλιώς μεταβαίνει ξανά στο βήμα 2.

Το πρώτο βήμα του αλγορίθμου FRM είναι η αρχικοποίηση του συστήματος με τους αρχικούς πληθυσμούς των μοριακών ειδών και η αρχικοποίηση του χρόνου προσομοίωσης. Στη συνέχεια επαναλαμβάνονται συνέχεια τα βήματα 2 έως 5 έως ότου ο χρόνος προσομοίωσης t γίνει ίσος ή μεγαλύτερος του προκαθορισμένου από το χρήστη χρόνου προσομοίωσης T_{sim} . Κατά την εκτέλεση των βημάτων 2 έως 5 πραγματοποιείται ο υπολογισμός όλων των τάσεων α_j των αντιδράσεων R_j και ο υπολογισμός όλων των χρονικών βημάτων τ_j που απαιτούνται για να ενεργοποιηθεί η κάθε αντίδραση R_j . Η αντίδραση R_μ με το μικρότερο δυνητικά χρόνο ενεργοποίησης τ_μ επιλέγεται και με βάση αυτή την αντίδραση ανανεώνεται η κατάσταση ($x = x + \nu_\mu$) και ο χρόνος προσομοίωσης του συστήματος ($t = t + \tau_\mu$). Η εκτέλεση των βημάτων 2 έως 5 ορίζεται ως ένας κύκλος αντιδράσεων (reaction cycle).

Μελετώντας τους δύο αλγορίθμους, Direct Method-SSA (DM-SSA) και FRM-SSA, προκύπτουν πλεονεκτήματα και μειονεκτήματα για κάθε αλγόριθμο. Ο αλγόριθμος DM-SSA απαιτεί τον υπολογισμό και την αποθήκευση όλων των τάσεων α_j σε ένα κύκλο αντιδράσεων ενώ θα πρέπει να υπολογιστεί και το συνολικό άθροισμα $\alpha_0(x)$ όλων των τάσεων κατά το δεύτερο βήμα. Αφού υπολογιστεί το συνολικό άθροισμα θα πρέπει να γίνουν σταδιακές αθροίσεις και συγκρίσεις κατά το πέμπτο βήμα του αλγορίθμου. Αντιθέτως στον αλγόριθμο FRM-SSA οι τάσεις των αντιδράσεων α_j δεν αθροίζονται ούτε αποθηκεύονται αλλά

χρησιμοποιούνται για τον υπολογισμό του χρόνου ενεργοποίησης τ_j της εκάστοτε αντίδρασης R_j . Ταυτόχρονα με τον υπολογισμό των τ_j χρόνων ενεργοποίησης είναι δυνατόν να εκτελείται και το βήμα 4 του αλγορίθμου που αφορά την εύρεση του μικρότερου τ_j χρόνου και την αντίδραση R_μ από την οποία προέκυψε. Το μόνο που απαιτείται να γίνει είναι να συγκρίνεται κάθε φορά ο τρέχων χρόνος τ_j με τον αμέσως προηγούμενο χρόνο τ_{j-1} που υπολογίστηκε. Επιπλέον, είναι σαφές από τα βήματα του αλγορίθμου FRM-SSA ότι δεν υπάρχει καμία συσχέτιση μεταξύ των αντιδράσεων R_j . Οι παραπάνω διαφοροποιήσεις συντελούν στο γεγονός ότι ο αλγόριθμος FRM-SSA μπορεί να παραλληλοποιηθεί σε μεγάλο βαθμό σε αντίθεση με τον αλγόριθμο DM-SSA. Ένα μειονέκτημα του FRM-SSA αλγορίθμου είναι ότι απαιτούνται m ψευδοτυχαίοι αριθμοί σε κάθε κύκλο αντιδράσεων σε σύγκριση με τους δύο ψευδοτυχαίους αριθμούς που απαιτούνται στον αλγόριθμο DM-SSA.

2.5 Ο αλγόριθμος NRM-SSA των Gibson και Bruck

Μια πολύ ενδιαφέρουσα παρατήρηση που μπορεί να γίνει στους αλγορίθμους του Gillespie είναι ότι συνήθως, η αντίδραση R_μ δεν επηρεάζει τις τάσεις α_j όλων των αντιδράσεων ενός βιομοντέλου. Κάτι τέτοιο σημαίνει ότι, δεν είναι απαραίτητο να υπολογίζονται, σε κάθε κύκλο αντιδράσεων, όλες οι τάσεις των αντιδράσεων παρά μόνο οι τάσεις των αντιδράσεων που επηρεάστηκαν από τη τελευταία R_μ αντίδραση. Οι Gibson και Bruck καταλήγοντας στο παραπάνω συμπέρασμα επινόησαν τον αλγόριθμο στοχαστικής προσομοίωσης Next Reaction Method (NRM-SSA) [37] ο οποίος εκμεταλλευόμενος αυτή τη παρατήρηση έχει πολυπλοκότητα εκτέλεσης $O(\log(m))$ σε σύγκριση με τον αλγόριθμο FRM-SSA που έχει πολυπλοκότητα εκτέλεσης $O(m)$.

Ο αλγόριθμος NRM-SSA διαφοροποιείται από τον αλγόριθμο FRM-SSA σε τρία σημεία. Πρώτον αλλάζει η έννοια του χρόνου, διότι ο FRM-SSA αλγόριθμος υπολογίζει την σχετική απόσταση του χρόνου της επόμενης αντίδρασης από την τρέχουσα χρονική στιγμή, σε αντίθεση με τον NRM-SSA αλγόριθμο που υπολογίζει την καθολική τιμή του χρόνου την οποία θα εκτελεστεί η επόμενη αντίδραση. Δεύτερον, όπως αναφέρεται και παραπάνω, οι Gibson και Bruck εισήγαγαν την έννοια της εξάρτησης ανάμεσα στις αντιδράσεις. Κάθε αντίδραση που εκτελείται δεν επηρεάζει τις τάσεις $\alpha_j(x)$ όλων των αντιδράσεων, αλλά μόνο των επηρεαζόμενων αντιδράσεων. Επομένως, απαιτείται η δημιουργία ενός κατευθυνόμενου γράφου εξαρτήσεων (Dependency Graph) όπου η ακμή από την αντίδραση με τον μικρότερο χρόνο προς άλλες αντιδράσεις υποδηλώνει ότι τουλάχιστον ένα αντιδρών ή ένα προϊόν της νικητήριας αντίδρασης συμμετέχει ως αντιδρών στις αντιδράσεις αυτές, τις "επηρεαζόμενες αντιδράσεις". Τρίτον και τέλος, δεδομένου ότι δεν χρειάζεται να υπολογιστούν όλες οι τάσεις $\alpha_j(x)$ και ο χρόνος τ_j των μη επηρεαζόμενων αντιδράσεων οι τιμές αυτές πρέπει να αποθηκευτούν. Με βάση αυτή τη δομή, σε κάθε κύκλο αντιδράσεων υπολογίζεται εκ νέου η τάση α_μ της αντίδρασης με τον μικρότερο χρόνο και ανανεώνονται με διαφορετικό μαθηματικό τύπο οι τάσεις $\alpha_j(x)$ των επηρεαζόμενων αντιδράσεων.

3. ΠΡΟΔΙΑΓΡΑΦΕΣ ΣΥΣΤΗΜΑΤΟΣ

Στην ενότητα που ακολουθεί θα παρουσιαστούν οι προδιαγραφές και ο τρόπος λειτουργίας του συστήματος που αναπτύχθηκε. Στα πλαίσια της παρούσας έρευνας αναπτύχθηκε ένα παράλληλο υπολογιστικό σύστημα υψηλών επιδόσεων για τη στοχαστική προσομοίωση μεγάλων βιομοριακών δικτύων με τη μέθοδο First Reaction Method-FRM. Το σύστημα απαρτίζεται από έναν υπολογιστή γενικού σκοπού (Host-PC – frontend) ο οποίος επικοινωνεί με το παράλληλο υπολογιστικό σύστημα υλικού FRM SoC (First Reaction Method System On Chip – backend) που αναπτύχθηκε για την εκτέλεση του αλγορίθμου FRM-SSA. Η περιγραφή του συστήματος έγινε σε γλώσσα περιγραφής υλικού VHDL και η υλοποίηση του έγινε σε ολοκληρωμένο κύκλωμα επαναπρογραμματιζόμενης λογικής (Field Programmable Gate Array - FPGA). Πιο συγκεκριμένα, για την ανάπτυξη του συστήματος χρησιμοποιήθηκε η FPGA αναπτυξιακή κάρτα υλικού (development board) KC705 της εταιρίας Xilinx [38]. Στο Host-PC βρίσκονται αποθηκευμένα τα δεδομένα των μοντέλων προς προσομοίωση ενώ είναι υπεύθυνο για κατάλληλη τροφοδότηση τους στο σύστημα FRM SoC. Επιπλέον, στο Host-PC αποθηκεύονται τα αποτελέσματα των προσομοιώσεων που εκτελούνται στο σύστημα FRM SoC.

Στη συνέχεια ακολουθούν οι προδιαγραφές της αρχιτεκτονικής FRM SoC, το πρωτόκολλο επικοινωνίας του συστήματος FRM SoC με το Host PC και τέλος η λειτουργία και η ροή εργασιών συνολικά του υπολογιστικού συστήματος.

3.1 Προδιαγραφές αρχιτεκτονικής FRM SoC

Για τη παραλληλοποίηση του FRM-SSA αλγορίθμου, σχεδιάστηκε μια πολυπύρηνη αρχιτεκτονική, 32-bit, ειδικού σκοπού πλήρους διοχέτευσης (fully pipelined). Η αρχιτεκτονική έχει τη μορφή παραμετρικού "μαλακού πυρήνα" (soft IP core) ώστε να μπορεί να διαμορφωθεί κατά το χρόνο της σύνθεσης με συγκεκριμένες τιμές παραμέτρων. Ως εκ τούτου, είναι δυνατόν να παραχθούν εύκολα με τη χρήση σύνθεσης υλικού διαφορετικές υλοποιήσεις SoCs για FPGAs ανάλογα με τις απαιτήσεις του χρήστη.

Η αρχιτεκτονική FRM SoC έχει τη δυνατότητα παράλληλης προσομοίωσης μεγάλων βιομοριακών δικτύων. Σύμφωνα με τα στατιστικά υλοποίησης της αναπτυξιακής κάρτας που χρησιμοποιήθηκε κατά το στάδιο της σχεδίασης, το σύστημα είναι σε θέση να προσομοιώσει βιομοντέλα που περιέχουν έως και 4K αντιδράσεις και μοριακά είδη. Οι αντιδράσεις που υποστηρίζονται από το σύστημα πρέπει να διέπονται από τους νόμους της κινητικής (mass action), έως τρίτης τάξης, και να είναι μη αντιστρέψιμες, ενώ λαμβάνονται υπόψη όλες οι πιθανές στοιχειομετρίες [39]. Η αντιστρεψιμότητα μιας αντίδρασης δε περιορίζει το σύστημα εφόσον αυτή μπορεί να απλοποιηθεί και να περιγραφεί μέσω απλούστερων στοιχειωδών μη αντιστρέψιμων αντιδράσεων.

Η αρχιτεκτονική του συστήματος υποστηρίζει πράξεις κινητής υποδιαστολής σύμφωνα με το πρότυπο IEEE 754 – Single Precision Floating Point Arithmetic για τον υπολογισμό των τάσεων των αντιδράσεων [40].

Όπως προαναφέρθηκε το σύστημα είναι παραμετρικό σύμφωνα με τα χαρακτηριστικά ενός βιολογικού δικτύου. Στη συνέχεια ακολουθούν τα χαρακτηριστικά που παραμετροποιούν την αρχιτεκτονική FRM SoC:

- Αριθμός μοριακών ειδών – n . Η παράμετρος n όπως και οι υπόλοιποι παράμετροι δίνονται ως generic μεταβλητές στο top-level αρχείο της υλοποίησης του συστήματος. Η παράμετρος καθορίζει το βάθος των μνημών όπου αποθηκεύονται τα μοριακά είδη.
- Αριθμός αντιδράσεων – m . Η παράμετρος m καθορίζει το βάθος της μνήμης όπου αποθηκεύονται οι αντιδράσεις του δικτύου.

Επιπλέον η αρχιτεκτονική που σχεδιάστηκε είναι παραμετρική και ως προς τον αριθμό των πόρων που διαθέτει. Η ευελιξία αυτή καθορίζεται επίσης από generic παραμέτρους κατά το χρόνο της σύνθεσης. Στη συνέχεια ακολουθούν οι παράμετροι:

- Αριθμός των FRM πυρήνων - C . Η παράμετρος C καθορίζει τον αριθμό των FRM πυρήνων που θα διαθέτει το σύστημα μετά τη σύνθεση.
- Αριθμός επεξεργαστικών μονάδων ανά πυρήνα – N . Η παράμετρος N καθορίζει τον αριθμό των επεξεργαστικών μονάδων που θα διαθέτει ο κάθε FRM πυρήνας μετά τη σύνθεση.

Οι παράμετροι C και N επίσης καθορίζουν και τους τρεις τρόπους λειτουργίας του FRM SoC. Κάθε πυρήνας του συστήματος είναι σε θέση να διεξάγει μια προσομοίωση ενός βιομοντέλου όπου οι αντιδράσεις του θα εκτελούνται παράλληλα στις διαθέσιμες επεξεργαστικές μονάδες (FRM Processing Units - FPU) του πυρήνα. Η λειτουργία αυτή περιγράφεται ως Single Simulation In Parallel – SSIP. Όπως προαναφέρθηκε, η παράμετρος C καθορίζει τον αριθμό των πυρήνων του συστήματος κάτι το οποίο δίνει τη δυνατότητα εκτέλεσης παράλληλων διαφορετικών και ανεξάρτητων προσομοιώσεων ως προς το βιομοντέλο που προσομοιώνεται και ως προς τις παραμέτρους του. Η λειτουργία αυτή περιγράφεται ως Multiple Simulations In Parallel – MSIP. Ο τρίτος τρόπος λειτουργίας του συστήματος είναι ο συνδυασμός των δύο παραπάνω οδηγώντας σε ένα υβριδικό σύστημα πολλών FRM πυρήνων όπου ο κάθε πυρήνας διαθέτει πολλές επεξεργαστικές μονάδες FPU. Όλες οι παραπάνω παράμετροι περιορίζονται μόνο από το FPGA που διαθέτει ο τελικός χρήστης και όχι από τις VHDL περιγραφές. Επιπλέον οι παράμετροι αυτοί καθορίζουν μόνο το μέγιστο των διαθέσιμων πόρων και ο τελικός χρήστης είναι σε θέση να διαλέξει ποιους από τους προαναφερθέντες πόρους θα χρησιμοποιήσει κατά τα πειράματα που θα εκτελέσει. Για τη καλύτερη κατανόηση των παραπάνω ακολουθούν τρία παραδείγματα συνδυασμών FRM πυρήνων και επεξεργαστικών μονάδων FPU.

- **FRM-SSA $C \times N$ SoC, παραδείγματα**
 1. **1x32** – SSIP τρόπος λειτουργίας, σύστημα ενός FRM πυρήνα οποίος διαθέτει 32 FPU
 2. **16x1** – MSIP τρόπος λειτουργίας, σύστημα 16 FRM πυρήνων όπου ο κάθε πυρήνας διαθέτει μια FPU
 3. **4x8** – Υβριδικός τρόπος λειτουργίας, σύστημα με 4 FRM πυρήνες όπου ο κάθε πυρήνας διαθέτει 8 FPU

3.2 Επικοινωνία Συστήματος

Προκειμένου να μπορέσει να εκτελεστεί μια προσομοίωση στο σύστημα απαιτούνται δεδομένα που σχετίζονται με το βιομοντέλο και αφορούν τις αρχικές τιμές των μοριακών ειδών, το σύνολο αντιδράσεων που περιγράφει ένα βιομοντέλο, η στοιχειομετρία των αντιδράσεων, η σταθερά του ρυθμού αντίδρασης, η τάξη της αντίδρασης

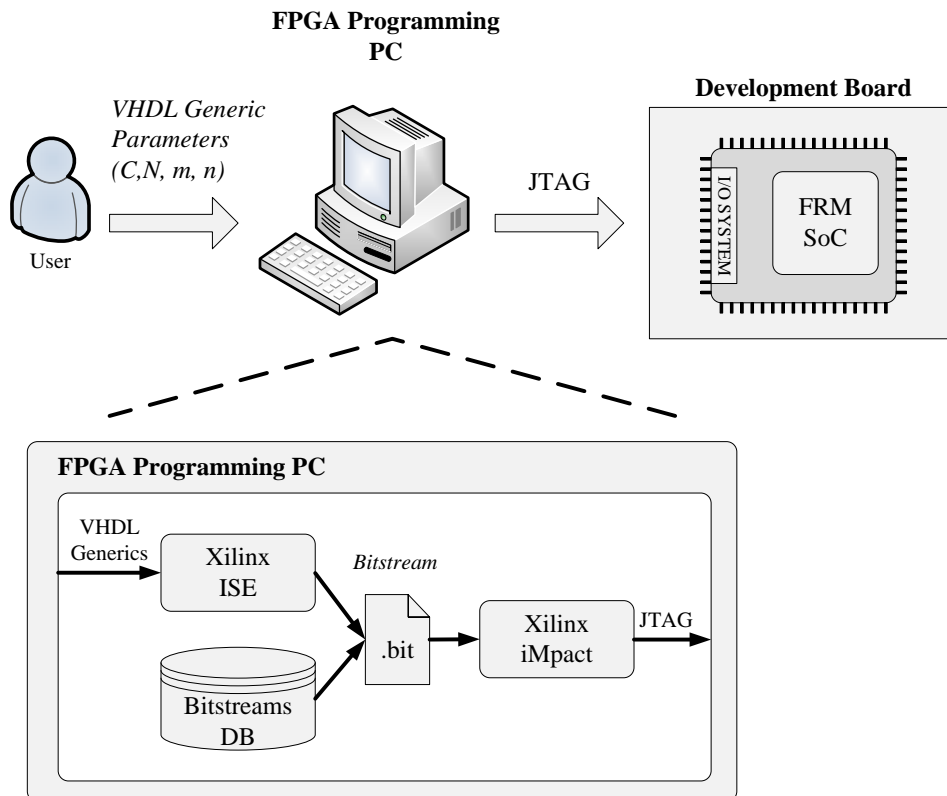
κ.α. Από την αρχή της σχεδίασης του συστήματος ήταν ξεκάθαρη απαίτηση όλα τα δεδομένα που χρειάζεται το σύστημα για να ξεκινήσει μια προσομοίωση να μην είναι αποθηκευμένα ήδη σε αυτό (κατά το χρόνο της σύνθεσης χρησιμοποιώντας μνήμες ROM) αλλά να μπορούν να αποθηκεύονται δυναμικά σε μνήμες RAM. Αυτή η απαίτηση δίνει τη δυνατότητα στο χρήστη να μπορεί να δοκιμάσει διαφορετικά σενάρια προσομοιώσεων στο FRM SoC χωρίς να απαιτείται ξανά η σύνθεση του συστήματος. Επιπλέον, το σύστημα εκτός από τα δεδομένα που χρειάζεται για να εκτελέσει μια προσομοίωση παράγει και αποτελέσματα που σχετίζονται με τις μεταβολές των μοριακών πληθυσμών κατά τη διάρκεια της προσομοίωσης. Όπως προαναφέρθηκε, όλα τα δεδομένα της προσομοίωσης αποθηκεύονται στο Host-PC. Για την επικοινωνία του Host-PC με το σύστημα FRM SoC επιλέξαμε να χρησιμοποιήσουμε το πρωτόκολλο μεταφοράς δεδομένων PCIe, εφόσον κάτι τέτοιο επιτρέπεται και από τους διαθέσιμους πόρους που έχουμε.

Ο λόγος που επιλέξαμε το συγκεκριμένο πρωτόκολλο σχετίζεται με τον τεράστιο όγκο των δεδομένων που μπορεί να παράξει το FRM SoC ανάλογα με τις παραμέτρους της προσομοίωσης. Έπειτα από την θεωρητική ανάλυση που διεξήχθη κατά τη σχεδίαση, προέκυψε ότι η μεταφορά δεδομένων από το FRM SoC σύστημα στο Host-PC μπορεί να αποτελέσει το κυριότερο παράγοντα επιβράδυνσης της λειτουργίας του συστήματος. Η ανάλυση μας κατέληξε στο ότι προκειμένου να αποφευχθεί κάτι τέτοιο θα πρέπει να χρησιμοποιηθεί ένα πραγματικά ταχύρρυθμο πρωτόκολλο επικοινωνίας. Εξετάζοντας τα πρωτόκολλα επικοινωνίας που είχαμε στη διάθεση μας (RS-232 [41], IEEE 802.3 - Ethernet [42], PCIe [43]) και έχοντας ως γνώμονα τη ταχύρρυθμη μεταφορά δεδομένων καταλήξαμε να χρησιμοποιήσουμε το πρωτόκολλο επικοινωνίας PCIe.

Για να μπορέσουμε να χρησιμοποιήσουμε το PCIe πρωτόκολλο χρειάστηκε επιπλέον λογισμικό (PCIe Drivers) από τη πλευρά του Host PC και επιπλέον υλικό από τη πλευρά του FPGA. Τους πυρήνες του επιπλέον λογισμικού και υλικού του συστήματος επικοινωνίας τους είχαμε στη διάθεση μας από γνωστές βιβλιοθήκες υλικού/λογισμικού που σχετίζονται με την PCIe επικοινωνία. Ωστόσο οι πυρήνες αυτοί από μόνοι τους δε συντελούν ένα σύστημα επικοινωνίας. Για αυτό το λόγο, στα πλαίσια του έργου, χρειάστηκε να γίνει περαιτέρω ανάπτυξη των προαναφερθέντων πυρήνων προκειμένου να πλαισιωθούν κατάλληλα και τελικά να μας δώσουν ένα ολοκληρωμένο σύστημα επικοινωνίας PCIe.

3.3 Λειτουργία Συστήματος και Ροή Εργασιών

Η λειτουργία του υπολογιστικού συστήματος χωρίζεται σε τρεις φάσεις. Στη πρώτη ο φάση ο χρήστης διαμορφώνει τη μορφή της αρχιτεκτονικής μέσω των Generic παραμέτρων στο υψηλότερο (top-level) αρχείο σχεδίασης VHDL. Αφού καθορίσει τις Generic παραμέτρους της αρχιτεκτονικής μέσω του λογισμικού ISE [44] της εταιρείας Xilinx παράγει το δυαδικό αρχείο bitstream [45] το οποίο περιέχει τη περιγραφή της αρχιτεκτονικής σε επίπεδο κυκλώματος. Έπειτα αφού παραχθεί το bitstream ο χρήστης προγραμματίζει το FPGA μέσω του λογισμικού iMpack [46], επίσης της εταιρείας Xilinx. Η μεταφορά του bitstream στο FPGA γίνεται μέσω του πρωτοκόλλου επικοινωνίας JTAG [47]. Εδώ θα πρέπει να αναφερθεί ότι σκοπός της παρούσας έρευνας είναι τελικά να δημιουργηθεί μια βιβλιοθήκη από bitstreams προκειμένου το βήμα της παραγωγής ενός bitstream να αποφεύγεται καθώς αποτελεί μια χρονοβόρα διαδικασία. Στόχος είναι να απαιτείται να δημιουργηθεί ένα νέο bitstream σε εξαιρετικές περιπτώσεις όπου οι ανάγκες κάποιου χρήστη εξειδικεύονται τόσο που επιβάλουν το βήμα αυτό. Στη συνέχεια ακολουθεί ένα σχηματικό που οπτικοποιεί τις προαναφερθείσες διαδικασίες.



Εικόνα 1: Προγραμματισμός του FPGA με την αρχιτεκτονική FRM SoC

Με την ολοκλήρωση της φάσης του προγραμματισμού του FPGA ο χρήστης μπορεί να προχωρήσει στη δεύτερη φάση. Αυτή η φάση περιλαμβάνει το προγραμματισμό του FRM SoC συστήματος με το επιθυμητό βιομοντέλο και την εκτέλεση του πειράματος. Σε αυτή τη φάση ο χρήστης αποθηκεύει στο Host-PC το βιομοντέλο που επιθυμεί να προσομοιώσει σε SBML κωδικοποίηση. Το SBML αρχείο περιέχει το σύνολο αντιδράσεων που περιγράφουν το βιολογικό μοντέλο καθώς και τα μοριακά είδη και τους αρχικούς πληθυσμούς τους. Τα δεδομένα αυτά πρέπει να μετατραπούν σε κατάλληλες δομές δεδομένων που αναγνωρίζει το FRM SoC σε δυαδική μορφή. Τη μετατροπή αυτή την αναλαμβάνει ένα εργαλείο λογισμικού που αναπτύχθηκε στα πλαίσια του έργου και ονομάζεται SBML Biomodel Parser. Εκτός από τις αντιδράσεις του βιομοντέλου και τα μοριακά είδη θα πρέπει να ρυθμιστούν και οι παράμετροι που σχετίζονται με το χρόνο προσομοίωσης ενός βιομοντέλου. Αυτές οι παράμετροι ρυθμίζονται από τον χρήστη μέσω του Αρχείου Ρυθμίσεων (configuration file). Αφού συμπληρωθεί και το αρχείο ρυθμίσεων τροφοδοτείται μαζί με το SBML αρχείο στο εργαλείο SBML Biomodel Parser προκειμένου να παραχθούν τα κατάλληλα δυαδικά αρχεία. Οι ρυθμίσεις προσομοίωσης που καλείται να συμπληρώσει ο χρήστης είναι οι ακόλουθες:

1. T_{sim} (s): χρόνος ολοκλήρωσης της προσομοίωσης σε δευτερόλεπτα
2. T_{sam} (s): περίοδος δειγματοληψίας των αποτελεσμάτων σε δευτερόλεπτα.
3. **Επανάληψεις Προσομοίωσης (Repetitions - R)**: αριθμός των επαναλήψεων που θα εκτελεστεί το πείραμα
4. **Μέγιστος Αριθμός Εσωτερικών Βημάτων (Max Internal Steps)**: μέγιστος αριθμό κύκλων αντιδράσεων (reaction cycles) που θα μεσολαβήσουν κατά τη διάρκεια μιας περιόδου δειγματοληψίας προτού τερματιστεί βίαια μια επανάληψη πειράματος.

5. **Αριθμός Επεξεργαστικών Μονάδων FPU - N :** η παράμετρος N θα πρέπει να δηλωθεί και στο αρχείο ρυθμίσεων καθώς είναι απαραίτητη από το εργαλείο SBML Biomodel Parser προκειμένου να δημιουργήσει τις κατάλληλες δομές δεδομένων που χρειάζεται το FRM SoC για να ξεκινήσει μια προσομοίωση.

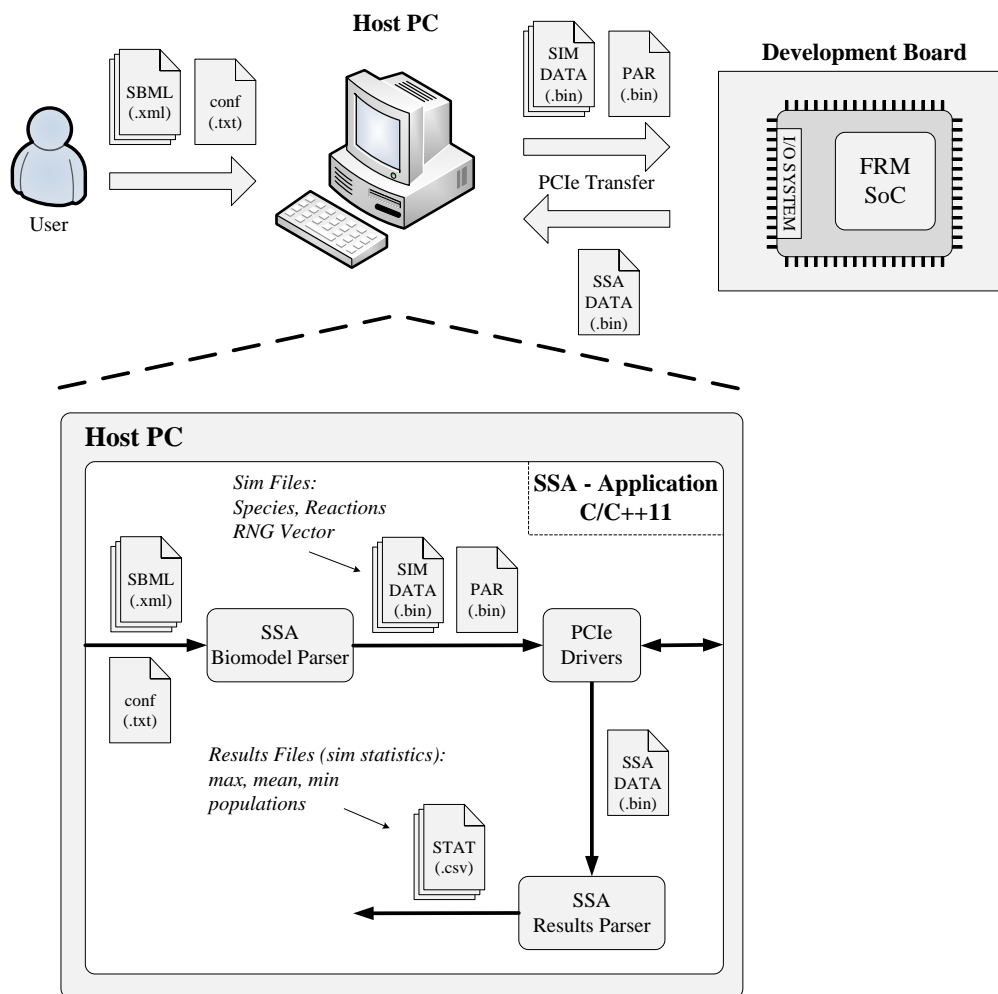
Η παράμετρος T_{sam} ουσιαστικά καθορίζει τις χρονικές στιγμές του χρόνου T_{sim} κατά τις οποίες θα καταγράφονται στο Host-PC οι τιμές των μοριακών πληθυσμών. Όπως θα αναφερθεί και στην Ενότητα 4 ο μηχανισμός της δειγματοληψίας των αποτελεσμάτων βοηθάει στο περιορισμό των παραγόμενων δεδομένων. Ο μηχανισμός του *Max Internal Steps* δίνεται για τον έλεγχο του πραγματικού χρόνου προσομοίωσης ο οποίος μπορεί να αυξηθεί σημαντικά στη περίπτωση όπου το σύστημα για κάποιο λόγο αρχίζει να παράγει πολύ μικρά χρονικά βήματα τ_m . Στη περίπτωση που ενεργοποιηθεί ο μηχανισμός αυτός το σύστημα θα μεταβεί αυτόματα στην επόμενη επανάληψη προσομοίωσης.

Όπως αναφέρθηκε και στην Ενότητα 3.1 το σύστημα FRM SoC αποτελείται από πολλούς FRM πυρήνες όπου ο κάθε πυρήνας μπορεί να εκτελέσει παράλληλες προσομοιώσεις βιομοντέλων, τελείως ανεξάρτητες μεταξύ τους. Για το λόγο αυτό το Αρχείο Ρυθμίσεων συμπληρώνεται κατά γραμμές. Κάθε γραμμή του αρχείου αποτελείται από τις πέντε παραπάνω παραμέτρους και καθορίζει τη λειτουργία ενός FRM πυρήνα. Συνεπώς, οι γραμμές του αρχείου ρυθμίσεων υποδηλώνουν και τον αριθμό των FRM πυρήνων που θα χρησιμοποιηθούν κατά την εκτέλεση του πειράματος. Στο σημείο αυτό θα πρέπει να αναφερθεί ότι δεν υπάρχει δέσμευση να χρησιμοποιηθούν όλοι οι FRM πυρήνες που έχει στη διάθεση του FRM SoC σύστημα. Αυτό συνεπάγεται ότι με μια διαμόρφωση $C \times N$ μπορούν να εκτελεστούν πολλά σενάρια προσομοιώσεων χωρίς να απαιτείται ξανά ο προγραμματισμός του FPGA με τη προϋπόθεση ότι τα σενάρια αυτά δεν υπερβαίνουν το μέγιστο των διαθέσιμων πόρων.

Στην Εικόνα 2 που ακολουθεί δίνεται ένα σχηματικό που δείχνει τη ροή εργασιών στο Host PC με σκοπό τη καλύτερη κατανόηση των όσων έχουν αναφερθεί μέχρι το σημείο αυτό. Αφού δοθούν τα κατάλληλα αρχεία και παράμετροι στο εργαλείο SBML Biomodel Parser παράγονται τα δυαδικά αρχεία με τις δομές δεδομένων που σχετίζονται με τις αντιδράσεις του βιομοντέλου. Στην Ενότητα 2 αναφέρθηκε ότι ο αλγόριθμος FRM-SSA χρειάζεται για κάθε αντίδραση R_j που εξετάζει και έναν ψευδοτυχαίο αριθμό r_j . Οι ψευδοτυχαίοι αυτοί αριθμοί παράγονται από γεννήτριες ψευδοτυχαίων αριθμών (pseudorandom number generator - RNG [48]) οι οποίες χρειάζεται να αρχικοποιηθούν κατάλληλα προτού ξεκινήσει η λειτουργία τους. Για την αρχικοποίηση της παράγονται κατάλληλα διανύσματα από το εργαλείο SBML Biomodel Parser τα οποία και στέλνονται από τη μέσω της PCIe επικοινωνίας στο FRM SoC. Όλα αυτά τα δυαδικά αρχεία στην Εικόνα 2 συμβολίζονται ως Sim Data (simulation data files). Εκτός από τα προαναφερθέντα αρχεία που σχετίζονται με το βιομοντέλο στο FRM SoC θα πρέπει να σταλούν και οι παράμετροι που σχετίζονται με τη προσομοίωση. Οι παράμετροι αυτές γράφονται από σε κατάλληλη δυαδική μορφή σε ένα αρχείο το οποίο στέλνεται στο FRM SoC μέσω της PCIe επικοινωνίας. Το αρχείο αυτό στην Εικόνα 2 συμβολίζεται ως PAR και υπεύθυνο για τη δημιουργία του είναι το εργαλείο SBML Biomodel Parser. Το αρχείο PAR αποτελείται από σύνολα λέξεων των 32-bit όπου το κάθε σύνολο αντιστοιχεί σε έναν FRM πυρήνα που πρόκειται να χρησιμοποιηθεί κατά τη προσομοίωση. Κάθε σύνολο αποτελείται από 7 λέξεις που κωδικοποιούν τις παραμέτρους 1 έως 4 του Αρχείου Ρυθμίσεων συν τις παρακάτω παραμέτρους:

- Συνολικό αριθμό μοριακών ειδών ενός βιομοντέλου.
- Συνολικό αριθμό αντιδράσεων ενός βιομοντέλου και αριθμό αντιδράσεων που αντιστοιχούν σε μια επεξεργαστική μονάδα FPU.

- Συνολικός αριθμός των ενεργών FRM πυρήνων και αναγνωριστικά (IDs) κάθε FRM πυρήνα.
- Τέλος στέλνεται και μια παράμετρος για τον έλεγχο του ρυθμού μετάδοσης δεδομένων από το FRM SoC στο Host PC.



Εικόνα 2: Ροή εργασιών προσομοίωσης στο υπολογιστικό σύστημα Host PC και FRM SoC

Με την ολοκλήρωση της προσομοίωσης ξεκινάει η τρίτη φάση του συστήματος η οποία περιλαμβάνει την ανάλυση των αποτελεσμάτων. Τα αποτελέσματα που παράγονται από τους FRM πυρήνες μεταφέρονται μέσω του PCIe πρωτοκόλλου στο Host PC όπου και αποθηκεύονται σε ένα δυαδικό αρχείο. Στην Εικόνα 2 το αρχείο αυτό συμβολίζεται ως SSA DATA. Με το τέλος της προσομοίωσης το αρχείο SSA DATA τροφοδοτείται σε ένα εργαλείο λογισμικού που αναπτύχθηκε στα πλαίσια του έργου και είναι υπεύθυνο για την ανάλυση των δεδομένων. Στην Εικόνα 2 το εργαλείο αυτό συμβολίζεται ως SSA Results Parser. Το εργαλείο SSA Results Parser αρχικά ξεχωρίζει τα δεδομένα που παρήγαγε ο κάθε FRM πυρήνας αφού είναι δυνατόν κάθε πυρήνας να εκτελεί ένα τελείως διαφορετικό πείραμα. Έπειτα για κάθε πείραμα που διεξήχθη το εργαλείο SSA Results Parser υπολογίζει μέσες τιμές, ελάχιστα και μέγιστα για όλους τους μοριακούς πληθυσμούς, εφόσον έχουν εκτελεστεί περισσότερες από μια επαναλήψεις. Τα στατιστικά αυτά μαζί με τα αποτελέσματα κάθε επανάληψης

κάθε FRM πυρήνα αποθηκεύονται σε CSV αρχεία [49] προκειμένου να μελετηθούν περαιτέρω από τον χρήστη. Επιπλέον το εργαλείο SSA Results Parser δημιουργεί CSV αρχεία με στατιστικά που σχετίζονται με τη προσομοίωση όπως: τον συνολικό αριθμό κύκλων αντιδράσεων που εκτελέστηκαν, τον συνολικό αριθμό περιόδων δειγματοληψίας που καταγράφηκαν καθώς και τον χρόνο προσομοίωσης t_{sim} που μεσολάβησε για να ολοκληρωθεί η κάθε επανάληψη.

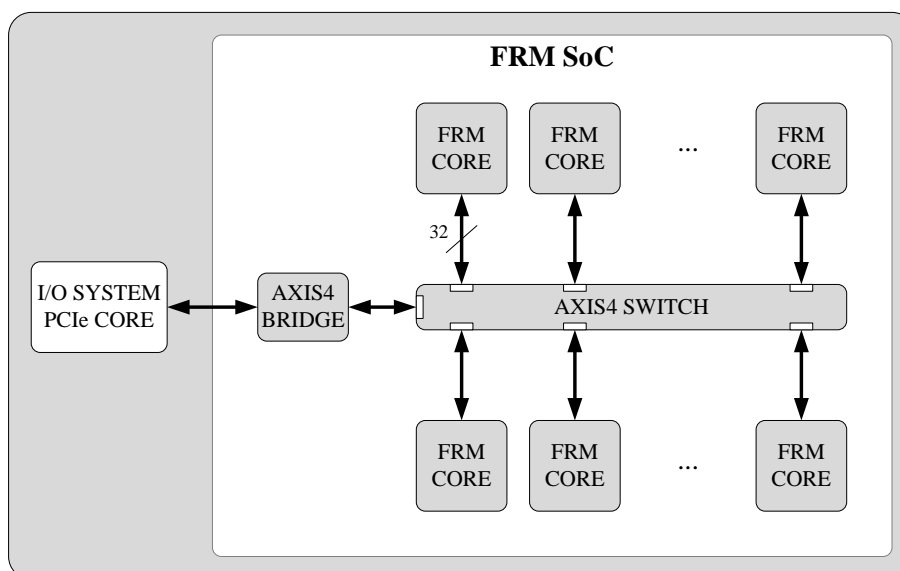


4. ΑΡΧΙΤΕΚΤΟΝΙΚΗ FRM ΣΥΣΤΗΜΑΤΟΣ ΣΕ ΨΗΦΙΔΑ

Στις προηγούμενες ενότητες, παρουσιάστηκαν οι αλγόριθμοι στοχαστικής προσομοίωσης και οι προδιαγραφές του συστήματος σε ψηφίδα (Systems on Chip) που αναπτύχθηκε. Στην Ενότητα που ακολουθεί θα γίνει εκτενής παρουσίαση της αρχιτεκτονικής FRM SoC και του FRM πυρήνα, ενώ στην επόμενη ενότητα θα παρουσιαστεί η αρχιτεκτονική της FRM επεξεργαστικής μονάδας. Η παρουσίαση του συστήματος θα ξεκινήσει από το υψηλότερο επίπεδο και σταδιακά θα παρουσιαστούν όλες οι μονάδες του συστήματος χρησιμοποιώντας όπου είναι απαραίτητο σχηματικά και διαγράμματα ροής δεδομένων.

4.1 Επίπεδο Συστήματος - Δίκτυο FRM Πυρήνων

Το υψηλότερο επίπεδο σχεδίασης (top-level) του FRM SoC αποτελείται από ανεξάρτητους πυρήνες FRM (FRM Core) οι οποίοι επικοινωνούν μέσω ενός διανομέα (SWITCH) με το PCIe σύστημα επικοινωνίας. Το σύστημα μπορεί να θεωρηθεί ως ένα δίκτυο (Network On Chip) από FRM πυρήνες με τοπολογία αστέρα. Στη συνέχεια ακολουθεί σχηματικό αρχιτεκτονικής δικτύου.



Εικόνα 3: FRM-SSA SoC, αρχιτεκτονική συστήματος

Οι διεπαφές των μονάδων του δικτύου που υλοποιήθηκε ακολουθούν πλήρως τις προδιαγραφές του πρωτοκόλλου AXI4 του διαύλου AMBA [50] [51]. Το AXI4 υποστηρίζει τις τρεις ακόλουθες εκδόσεις: AXI4, AXI4-Lite και AXI4-Stream Interface (AXIS4). Οι δύο πρώτες εκδόσεις στηρίζονται σε δομές δεδομένων μνήμης για την επικοινωνία των μονάδων ενώ η έκδοση AXIS4 στην απευθείας μεταφορά δεδομένων μεταξύ των μονάδων κάτι το οποίο και το καθιστά ταχύτερο πρωτόκολλο σε σύγκριση με τα δύο πρώτα. Δεδομένου της απαίτησης για ταχύρυθμη μεταφορά δεδομένων από και προς το SWITCH επιλέξαμε να χρησιμοποιήσουμε το AXIS4 ως πρωτόκολλο επικοινωνίας του δικτύου, εφόσον βάση των προδιαγραφών του συστήματος ήταν εφικτό. Το πρωτόκολλο επικοινωνίας AXI4 είναι το πιο δημοφιλές και διαδεδομένο πρωτόκολλο επικοινωνίας για τη διασύνδεση IP cores σε SoCs για FPGA υλοποιήσεις και όχι μόνο. Αξίζει να αναφερθεί ότι όλες οι

διεπαφές των νέων εκδόσεων IP cores της βιβλιοθήκης Core Generation της Xilinx χρησιμοποιούν αποκλειστικά το AXIS4 πρωτόκολλο επικοινωνίας. Επιπλέον η επιλογή ενός στάνταρτ και αξιόπιστου προτύπου όπως το AXIS4 έχει δύο ακόμη πλεονεκτήματα πέρα της ταχύτητας. Το πρώτο πλεονέκτημα αφορά το γεγονός ότι οι FRM πυρήνες μπορούν να προσαρμοστούν σε οποιοδήποτε σύστημα που υποστηρίζει το AXIS4 πρωτόκολλο. Το δεύτερο πλεονέκτημα που προκύπτει από τη σχεδίαση του δικτύου με αυτόν τον τρόπο είναι ότι μπορούν να προστεθούν μελλοντικά μονάδες (Microcontrolles, IP cores) που θα προσφέρουν νέες δυνατότητες στο σύστημα χωρίς να χρειάζεται να γίνει καμία τροποποίηση, εφόσον οι διεπαφές αυτών των μονάδων υποστηρίζουν το AXIS4.

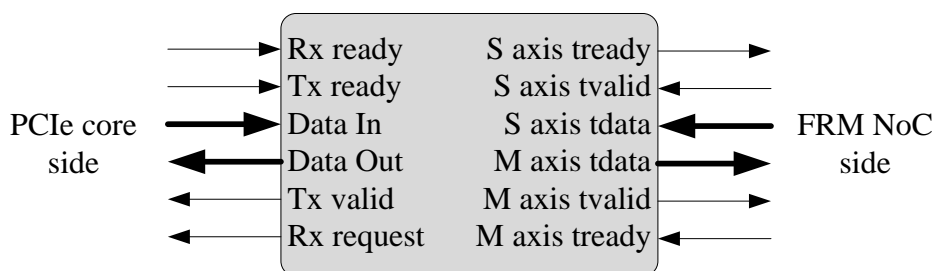
4.2 AXIS4 SWITCH – AXIS4 BRIDGE

Στην Εικόνα 4.1 φαίνεται η αρχιτεκτονική του SoC η οποία αποτελείται από πολλούς ανεξάρτητους FRM πυρήνες. Ο αριθμός των FRM πυρήνων εξαρτάται, όπως αναφέρθηκε και Στην Ενότητα 3, από τη Generic μεταβλητή *C* η οποία δηλώνεται στο υψηλότερο (top-level) αρχείο της υλοποίησης. Οι πυρήνες συνδέονται με το σύστημα επικοινωνίας (PCIe core) μέσω ενός SWITCH. Επίσης, ο αριθμός των υποδοχών του SWITCH εξαρτάται από τη Generic παράμετρο *C*. Οι FRM πυρήνες συνδέονται στους υποδοχείς μέσω 32-bit διαύλων ανά κατεύθυνση υποστηρίζοντας έτσι την αμφίδρομη επικοινωνία με τον διανομέα. Η κύρια αρμοδιότητα του SWITCH είναι ο χρονικός διαμοιρασμός (Time-division multiplexing) του καναλιού επικοινωνίας στους FRM πυρήνες.

Ο αλγόριθμος διαμοιρασμού (Bus Arbiter Algorithm) του καναλιού που υλοποιεί το SWITCH είναι ο Round Robin. Ο συγκεκριμένος αλγόριθμος είναι η πρώτη επιλογή για την σχεδίαση διανομέων λόγω της απλότητας του και του μικρού κόστους υλοποίησης. Η VHDL υλοποίηση της μονάδας SWITCH έγινε με τη χρήση Finite States Machines Datapath (FSMD). Η λειτουργία του διανομέα συνοψίζεται σε δύο φάσεις, αποκωδικοποίηση (datapath decoding) και διαμοιρασμός του καναλιού επικοινωνίας (Bus Arbiter). Κατά το στάδιο της αποκωδικοποίησης, ο διανομέας αποφασίζει σε ποιόν FRM πυρήνα θα πρέπει να στείλει τα δεδομένα που καταφθάνουν στο FRM SoC. Εφόσον το δίκτυο στηρίζεται στο AXI πρωτόκολλο, το είδος της επικοινωνίας είναι τύπου Master-Slave, όπου κατά το στάδιο της αποκωδικοποίησης ο διανομέας έχει το ρόλο του Master και οι FRM πυρήνες το ρόλο του Slave. Κατά το στάδιο του διαμοιρασμού, ο διανομέας δέχεται τα δεδομένα των FRM πυρήνων και σύμφωνα με τον αλγόριθμο Round Robin αποφασίζει σε ποιον FRM πυρήνα θα διαθέσει το κανάλι επικοινωνίας. Αν υπάρχουν πολλοί πυρήνες που ταυτόχρονα θέλουν να μεταδώσουν δεδομένα θα πρέπει να περιμένουν να έρθει η σειρά τους. Στο στάδιο του διαμοιρασμού, σύμφωνα και με το AXI πρωτόκολλο, το ρόλο του Master τον έχουν οι FRM πυρήνες και το ρόλο του Slave ο διανομέας.

Η διεπαφή του I/O συστήματος (PCIe core) δεν στηρίζεται στο AXIS4 πρωτόκολλο. Η βασική διαφοροποίηση της διεπαφής του I/O συστήματος είναι ότι και κατά τις δύο κατευθύνσεις μετάδοσης δεδομένων, από και προς το FRM SoC, το ρόλο του Slave τον έχει το I/O σύστημα. Για τη σύνδεση του FRM δικτύου με τον PCIe πυρήνα υλοποιήθηκε μια μονάδα γεφύρωσης (bridge). Στην Εικόνα 3 η μονάδα αυτή αναφέρεται ως AXIS4 BRIDGE και συνδέεται σε μια από τις υποδοχές της μονάδας SWITCH. Στην Εικόνα 4 δίνεται σχηματικό με τα σήματα ελέγχου και τις αρτηρίες δεδομένων της μονάδας AXIS4 BRIDGE.

AXIS4 BRIDGE



Εικόνα 4: Μπλοκ διάγραμμα της μονάδας AXIS4 BRIDGE

Στους δύο πίνακες που ακολουθούν γίνεται η σύντομη περιγραφή των σημάτων της Εικόνας 4. Στον πρώτο πίνακα περιγράφονται τα σήματα της διεπαφής I/O System – AXIS4 BRIDGE και στον δεύτερο πίνακα περιγράφονται τα σήματα της διεπαφής AXIS4 BRIDGE – FRM SoC.

Πίνακας 1: Περιγραφή των σημάτων του πρωτοκόλλου AXIS4

AXIS4 BRIDGE Interface – I/O System Side (PCIe core)	
Σήματα Ελέγχου/Διάυλοι	Περιγραφή
<i>Rx ready</i>	Το PCI core έχει διαθέσιμα δεδομένα προς λήψη
<i>Tx ready</i>	Το PCI core είναι σε θέση να δεχτεί δεδομένα για αποστολή
<i>Rx request</i>	Αίτημα προς το PCI core για λήψη δεδομένων
<i>Tx valid</i>	Το PCI core λαμβάνει έγκυρα δεδομένα
<i>Data In / Data Out</i>	Διάυλοι μεταφοράς δεδομένων από και προς το PCI core

AXIS4 BRIDGE Interface – Network Side (FRM SoC)	
Σήματα Ελέγχου/Διάυλοι	Περιγραφή
<i>S axis tready</i>	Το FRM SoC είναι σε θέση να δεχτεί δεδομένα
<i>S axis tvalid</i>	Τα δεδομένα που λαμβάνει το FRM SoC είναι έγκυρα
<i>S/M axis tdata</i>	Διάυλοι μεταφοράς δεδομένων από και προς το FRM SoC
<i>M axis tready</i>	Το FRM SoC έχει το δικαίωμα να στείλει δεδομένα
<i>M axis tvalid</i>	Το FRM SoC μεταδίδει έγκυρα δεδομένα

Προτού ξεκινήσει η περιγραφή του FRM πυρήνα και των υπομονάδων του είναι απαραίτητη η παρουσίαση των δομών δεδομένων που χρησιμοποιούνται για την οργάνωση και την αποθήκευση των δεδομένων στον πυρήνα. Στην επόμενη ενότητα παρουσιάζεται η μεθοδολογία και οι δομές που έχουν ακολουθηθεί για την επεξεργασία των δεδομένων ενός βιομοντέλου.

4.3 Δομές Δεδομένων του FRM πυρήνα

Το πρώτο και καθοριστικό βήμα της σχεδίασης του FRM πυρήνα ήταν η μελέτη των δομών δεδομένων που θα χρησιμοποιηθούν για την οργάνωση και την αποθήκευση των δεδομένων ενός βιομοντέλου. Όπως αναφέρθηκε και στην Ενότητα 2.1 τα βιομοντέλα αποτελούνται από αντιδράσεις που μοντελοποιούν τη συμπεριφορά του βιολογικού συστήματος. Σε ένα σύνολο αντιδράσεων, ένα μοριακό είδος S_i μπορεί να συμμετέχει σε πολλές αντιδράσεις. Προκειμένου να αποφευχθούν οι διπλοεγγραφές μοριακών πληθυσμών S_i στο σύνολο αντιδράσεων μπορούν να χρησιμοποιηθούν δείκτες προς τους πληθυσμούς των μοριακών ειδών S_i . Το κόστος των διπλοεγγραφών έχει δύο σκέλη. Το πρώτο σκέλος σχετίζεται με το ότι κάθε φορά που μεταβάλλεται ένας S_i πληθυσμός θα πρέπει να ανανεωθούν όλες οι αντιδράσεις στις οποίες συμμετέχει. Το δεύτερο σκέλος σχετίζεται με τον όγκο της αποθήκευσης δεδομένων. Η αναπαράσταση των S_i πληθυσμών γίνεται με αριθμητική κινητής υποδιαστολής απλής ακρίβειας. Κάτι τέτοιο σημαίνει ότι απαιτούνται 32-bit για την αποθήκευση κάθε S_i πληθυσμού σε αντίθεση με την αποθήκευση των δεικτών όπου χρησιμοποιείται ακέραια αριθμητική με εύρος $\text{Log}_2(n)$ -bit όπου η παράμετρος n σπάνια αναμένεται να είναι μεγαλύτερη από τη τιμή 2^{16} . Συνεπώς, χρησιμοποιώντας δείκτες προς τους μοριακούς πληθυσμούς σε μια αντίδραση, το σύνολο των αντιδράσεων δεν χρειάζεται καμία ανανέωση κατά τη διάρκεια μιας προσομοίωσης ενώ απαιτείται λιγότερος αποθηκευτικός χώρος για την καταχώρηση των αντιδράσεων.

Ακολουθώντας αυτή τη στρατηγική μπορούμε να διαχωρίσουμε τα δεδομένα ενός βιομοντέλου που αποθηκεύονται σε έναν FRM πυρήνα σε δύο κατηγορίες. Η πρώτη περιλαμβάνει δεδομένα που μεταβάλλονται κατά τη διάρκεια εκτέλεσης ενός πειράματος. Τα δεδομένα αυτά είναι οι πληθυσμοί X_i των μοριακών ειδών και αποθηκεύονται σε έναν δυσδιάστατο πίνακα (Species Table - ST) $n \times 32$ -bit. Η δεύτερη κατηγορία περιλαμβάνει δεδομένα που δεν μεταβάλλονται κατά τη διάρκεια εκτέλεσης ενός πειράματος. Τα δεδομένα αυτά είναι τα δεδομένα μιας αντίδρασης τα οποία επίσης αποθηκεύονται σε έναν δυσδιάστατο πίνακα (Reaction Table - RT). Συνεπώς, κάθε γραμμή του πίνακα RT έχει αποθηκευμένη μια αντίδραση με τα εξής πεδία:

- Οκτώ πεδία για την αποθήκευση των δεικτών προς τα μοριακά είδη που συμμετέχουν σε μια αντίδραση, Species Pointers – SP_i. Τα οχτώ πεδία χρησιμοποιούνται για την αποθήκευση 3 δεικτών προς τους μοριακούς πληθυσμούς των αντιδρώντων (Reaction Pointers - RP_i) της αντίδρασης ενώ τα υπόλοιπα 5 πεδία για την αποθήκευση 5 δεικτών προς τους μοριακούς πληθυσμούς των προϊόντων (Product Pointers - PP_i) της αντίδρασης.
- Ένα πεδίο για την αποθήκευση της σταθεράς K_i του ρυθμού αντίδρασης. Για τη σταθερά K_i χρησιμοποιείται αριθμητική κινητής υποδιαστολής απλής ακρίβειας, συνεπώς απαιτούνται 32 bit για την αποθήκευσή της.
- Ένα πεδίο για κωδικοποίηση του διανύσματος στοιχειομετρίας, Vector Table – VT. Το διάνυσμα της στοιχειομετρίας περιέχει 8 αριθμούς, ένα για κάθε πληθυσμό που συμμετέχει σε μια αντίδραση, των τεσσάρων bit. Το εύρος των τεσσάρων bit δίνει τη δυνατότητα για στοιχειομετρίες από 1 έως 15 ενώ σύμφωνα με την υπάρχουσα βιβλιογραφία συνηθίζονται στοιχειομετρίες από 1 έως 3. Η κωδική λέξη VT

αποκωδικοποιείται από τη μονάδα STOICHIO DECODER του FRM πυρήνα και παράγονται οχτώ λέξεις των 32-bit αριθμητικής κινητής υποδιαστολής απλής ακρίβειας που χρησιμοποιούνται στην ανανέωση των πληθυσμών της R_μ αντίδρασης.

- Ένα πεδίο για την αποθήκευση του είδους της αντίδρασης - FG. Οι τάξεις των αντιδράσεων ενός βιομοντέλου μπορούν να ποικίλουν. Επίσης μια αντίδραση κάποιας τάξης μπορεί να αποτελείται από ένα μοριακό είδος μόνο ή από διαφορετικούς συνδυασμούς μοριακών ειδών. Το πεδίο FG με εύρος 3-bit κωδικοποιεί όλους τους πιθανούς συνδυασμούς αντιδράσεων έως και τρίτης τάξης και χρησιμοποιείται στον υπολογισμό των τάσεων των αντιδράσεων.

Όλα τα παραπάνω πεδία είναι σταθερού εύρου εκτός από τα πεδία των οχτώ δεικτών SP_i τα οποία εξαρτώνται από το βάθος του πίνακα ST. Το εύρος μια λέξης του πίνακα RT μπορεί να υπολογιστεί από την ακόλουθη σχέση $RT\ WORD_w = 8 \times \text{Log}_2(N) + W$ όπου $W = K_w + VT_w + FG_w = 67\text{-bit}$. Το μέγεθος του πίνακα RT καθορίζεται από το γινόμενο $m \times RT\ WORD_w$. Στη συνέχεια ακολουθεί ένα παράδειγμα με τα περιεχόμενα των πινάκων RT και ST. Για το παράδειγμα χρησιμοποιήθηκε ξανά το βιομοντέλο LVS [52] του οποίου οι αντιδράσεις και οι παράμετροι του παρουσιάστηκαν στην Ενότητα 2.1.

Πίνακας 2: Η δομή των πινάκων RT και ST του FRM πυρήνα

Reaction Table - RT											
R_i	8 x $\text{Log}_2(n)$ 32-bit								32-bit	32-bit	3-bit
	Species Pointers - SP_i								K Rates	VT Code	Rctn Order
	Reactants				Products						
R_1	SP_1	—	—	SP_1	SP_1	—	—	—	K_1	VT_1	FG_1
R_2	SP_1	SP_2	—	SP_2	SP_2	—	—	—	K_2	VT_2	FG_2
R_3	SP_2	—	—	—	—	—	—	—	K_3	VT_3	FG_3
$R_{..}$	—	—	—	—	—	—	—	—	—	—	—
R_m	—	—	—	—	—	—	—	—	—	—	—

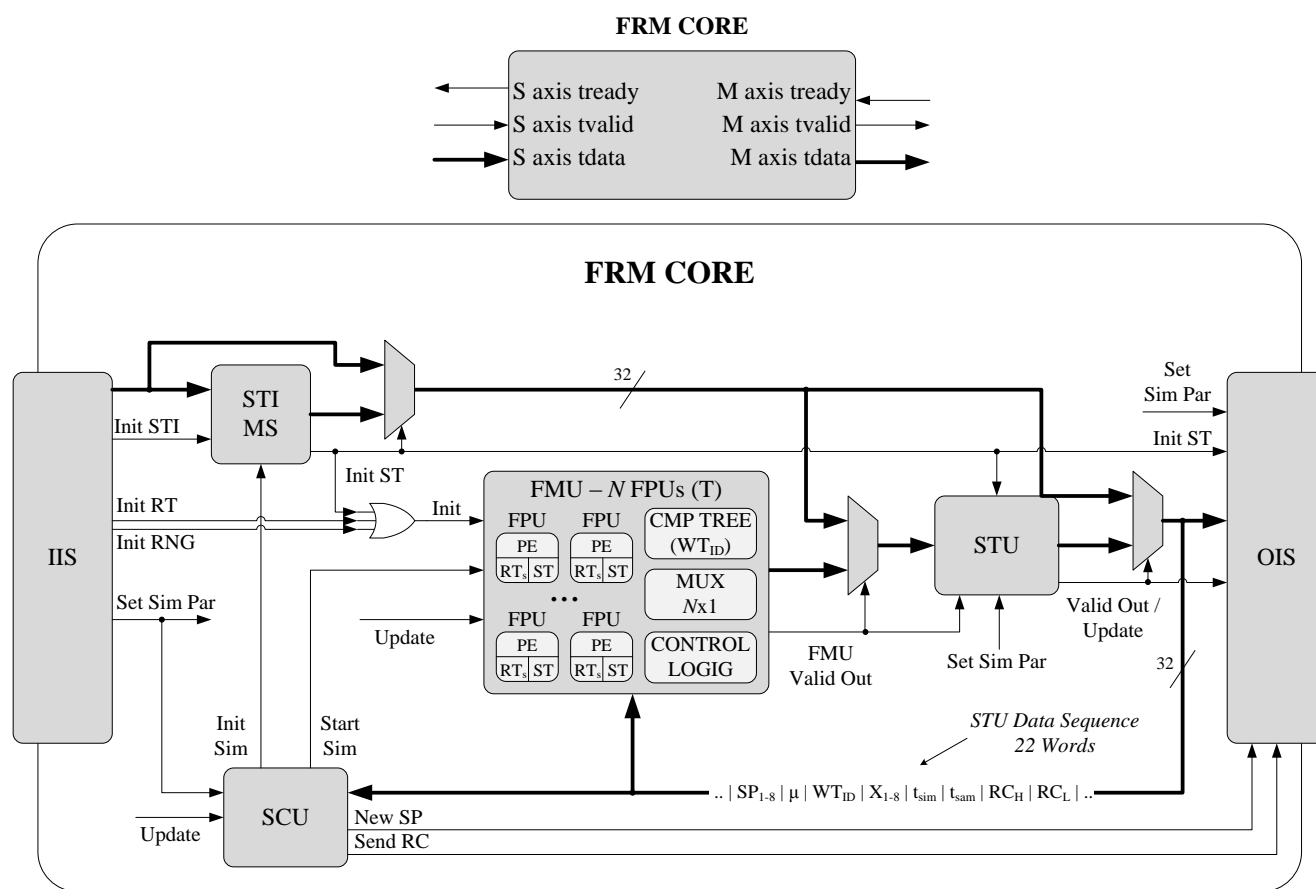
Species Table - ST	
Species Pointers	32-bit
	Particle Numbers
SP_0	0
SP_1	X_1
SP_2	X_2
$SP_{..}$	—
SP_n	—

Θα πρέπει να αναφερθεί ότι στον FRM πυρήνα η αποθήκευση των μοριακών πληθυσμών στον πίνακα ST ξεκινάει από τη θέση 1 και όχι από τη θέση 0. Η θέση 0 του πίνακα ST έχει πάντοτε αποθηκευμένη τη τιμή μηδέν και χρησιμοποιείται σε πράξεις όπου για διάφορους λόγους δε πρέπει να μεταβληθεί ένας μοριακός πληθυσμός παρότι συμμετέχει στην αντίδραση R_μ . Τέτοιες περιπτώσεις είναι μοριακά είδη που έχουν

χαρακτηριστεί ως πηγές (sources) του μοντέλου. Τα μοριακά είδη αυτά καθορίζονται από το SBML αρχείο και ο FRM πυρήνας σε αυτές τις περιπτώσεις χρησιμοποιεί τη τιμή της θέσης 0 του πίνακα ST κατά το στάδιο της ανανέωσης τους.

4.4 Αρχιτεκτονική FRM Πυρήνα

Μέχρι τώρα έχει γίνει παρουσίαση του FRM SoC σε επίπεδο δικτύου FRM πυρήνων, ενώ δόθηκαν σύντομες περιγραφές του πρωτοκόλλου AXIS4 και των μονάδων AXIS4 SWITCH και AXIS4 BRIDGE. Σε αυτήν και τις επόμενες ενότητες που ακολουθούν θα γίνει εκτενής παρουσίαση του FRM πυρήνα. Κάθε FRM πυρήνας είναι σε θέση να διεξάγει μια προσομοίωση ενός βιομοντέλου μιας ή περισσότερων στοχαστικών επαναλήψεων (repetitions), δίνοντας έτσι την δυνατότητα εκτέλεσης πολλών παράλληλων και τελείως ανεξάρτητων προσομοιώσεων στο FRM SoC σύστημα. Στη συνέχεια ακολουθεί σχηματικό αρχιτεκτονικής ενός FRM πυρήνα.



Εικόνα 5: Αρχιτεκτονική του FRM πυρήνα

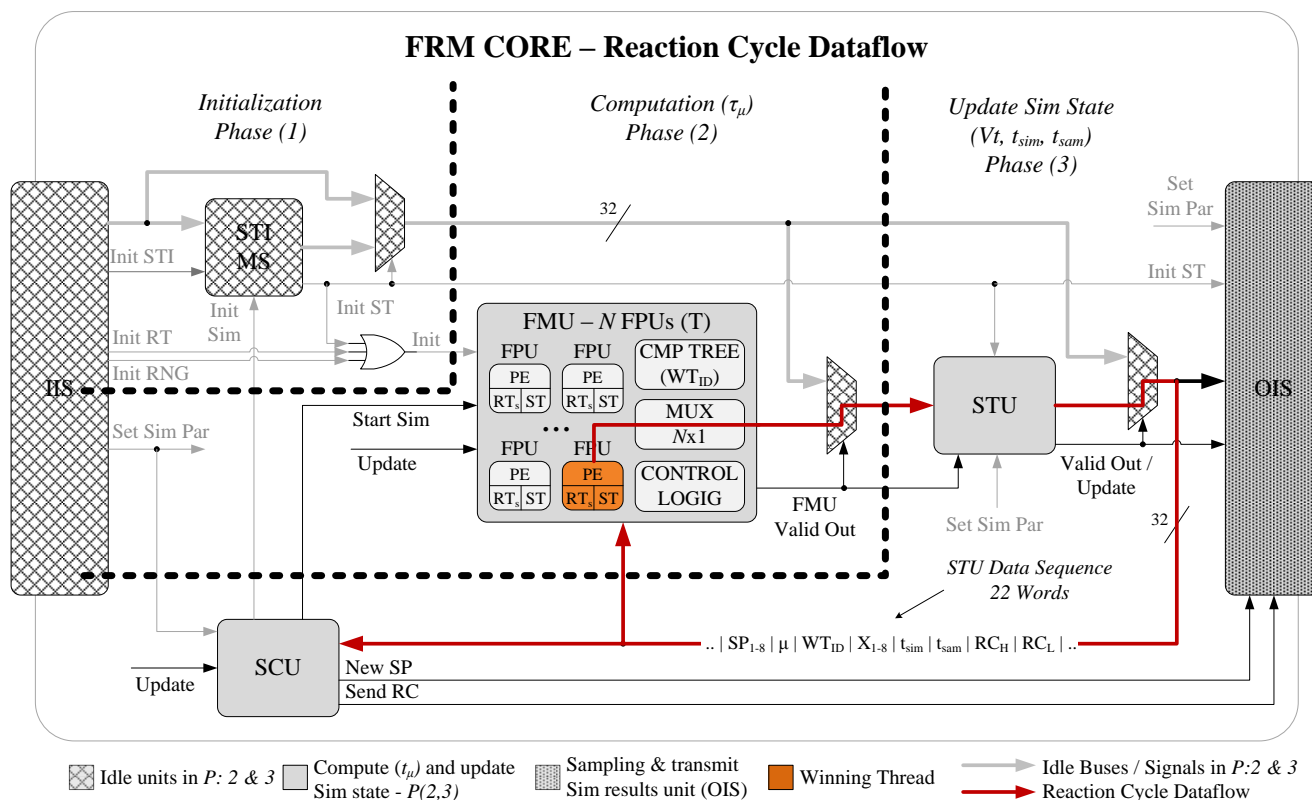
Το υψηλότερο επίπεδο του FRM πυρήνα αποτελείται από έξι μονάδες. Οι μονάδες IIS και OIS (Input και Output Interface Systems) είναι υπεύθυνες για την επικοινωνία του πυρήνα με το δίκτυο. Η μονάδα ελέγχου SCU (Simulation Control Unit) είναι υπεύθυνη για τον έλεγχο και την εξέλιξη της προσομοίωσης που

εκτελείται στο πυρήνα. Η μονάδα STI MS (ST Initialization Memory System) είναι υπεύθυνη για τη αποθήκευση των αρχικών πληθυσμών των μοριακών ειδών με σκοπό την αρχικοποίηση του συστήματος στην αρχή κάθε επανάληψης της προσομοίωσης. Η μονάδα FMU (FRM Multiprocessing Unit) είναι καρδιά του FRM πυρήνα και είναι υπεύθυνη για την εύρεση της επόμενης "νικητήριας αντίδρασης" R_μ που θα εκτελεστεί και του χρόνου τ_μ που απαιτείται για την "πυροδότηση" της. Τέλος, η μονάδα SPU (Special Purpose/Processing Unit) είναι υπεύθυνη για την ανανέωση των πληθυσμών των μοριακών ειδών και την ανανέωση των παραμέτρων της προσομοίωσης μετά το τέλος κάθε κύκλου αντιδράσεων.

Η FMU μονάδα αποτελείται από μια ή πολλές επεξεργαστικές μονάδες (FRM Processing Unit - FPU) με σκοπό την παραλληλοποίηση του FRM-SSA αλγορίθμου. Η παραλληλοποίηση του αλγορίθμου επιτυγχάνεται από το γεγονός ότι η κάθε FPU αποτελεί ένα ξεχωριστό νήμα υλικού (thread) που αναλαμβάνει να επεξεργαστεί ένα υποσύνολο (m/N) του συνολικού αριθμού αντιδράσεων του βιομοντέλου. Η FPU μονάδα αποτελείται από μνήμες για την αποθήκευση των πινάκων RT και ST που περιγράφηκαν στην Ενότητα 4.3 και από μία μονάδα (Processing Element - PE) για τον υπολογισμό του χρόνου τ_μ και την εύρεση της αντίδρασης (R_μ) από την οποία προέκυψε ο χρόνος αυτός. Στην παραπάνω εικόνα οι μνήμες αυτές συμβολίζονται ως RT section (m/N) – RTs και ως ST. Η FPU μονάδα σε κάθε κύκλο αντίδρασης (reaction cycle) αναλαμβάνει να υπολογίσει όλους τους χρόνους τ_j του υποσυνόλου (m/N) που της έχει ανατεθεί. Παράλληλα καθώς υπολογίζονται οι χρόνοι τ_j στην FPU μονάδα, συγκρατείται σε έναν καταχωρητή της μονάδας ο μικρότερος χρόνος που έχει υπολογιστεί προκειμένου στο τέλος της διαδικασίας να προκύψει ο χρόνος τ_μ . Για τον εντοπισμό του συνολικά μικρότερου χρόνου τ_j (τ_μ) όλων των αντιδράσεων χρησιμοποιείται ένα δυαδικό δέντρο συγκριτών (CMP TREE) και ένας πολυπλέκτης (MUX) $N \times 1$. Οι τοπικοί χρόνοι τ_μ που υπολογίστηκαν από κάθε FPU οδηγούνται στο δυαδικό δέντρο συγκριτών και από την σύγκριση τους προκύπτει το αναγνωριστικό του "νικητήριου νήματος" (Winning Thread ID - WT_{ID}) το οποίο υπολόγισε τον συνολικά μικρότερο χρόνο τ_j όλων των αντιδράσεων του πίνακα RT. Έπειτα μέσω του WT_{ID} και του πολυπλέκτη MUX $N \times 1$ τροφοδοτείται στην έξοδο της FMU το "νικητήριο νήμα" (Winning Thread - WT). Ως εκ τούτου στο τέλος κάθε κύκλου αντιδράσεων οδηγείται στην έξοδο της FMU μονάδας ο χρόνος τ_μ , ο δείκτης (μ) της "νικητήριας αντίδρασης" R_μ , το αναγνωριστικό WT_{ID} καθώς και τα δεδομένα της γραμμής του πίνακα RT που αντιστοιχούν στην αντίδραση R_μ .

Η λειτουργία του FRM πυρήνα διακρίνεται στις εξής τρεις φάσεις: αρχικοποίηση (*initialization*), υπολογισμός (computation) και ανανέωση (update). Ο FRM πυρήνας αρχικά μεταβαίνει στην φάση της αρχικοποίησης ενώ στη συνέχεια εναλλάσσεται μεταξύ των φάσεων υπολογισμού και ανανέωσης για την διεκπεραίωση ενός κύκλου αντιδράσεων (reaction cycle). Στην Εικόνα 6 δίνεται ξανά σχηματικό με την αρχιτεκτονική του FRM πυρήνα στο οποίο έχουν συμπεριληφθεί οι φάσεις αυτές. Επιπλέον στο σχηματικό της Εικόνας 6 δίνεται ένα παράδειγμα της ροής δεδομένων (dataflow) και των μονάδων που επιχειρούν κατά τη διάρκεια των φάσεων υπολογισμού και ανανέωσης.

Η πρώτη φάση (*Initialization - Phase 1*) χωρίζεται σε δύο στάδια. Το πρώτο στάδιο περιλαμβάνει την αποθήκευση των δεδομένων ενός βιομοντέλου (αντιδράσεις, αρχικοί πληθυσμοί μοριακών ειδών) στους πίνακες RT και STI. Επίσης κατά το στάδιο αυτό αποθηκεύονται στον FRM πυρήνα οι παράμετροι της προσομοίωσης που περιγράφηκαν στην Ενότητα 3.3. Το δεύτερο στάδιο περιλαμβάνει την αρχικοποίηση των πινάκων ST των FPU μονάδων με τα δεδομένα του πίνακα STI. Το στάδιο αυτό επαναλαμβάνεται στην αρχή κάθε επανάληψης της προσομοίωσης. Η μνήμη του πίνακα STI δρα ως κρυφή μνήμη (cache memory) και ο σκοπός της είναι να αποφευχθεί η συνεχής επικοινωνία με το Host PC για την αρχικοποίηση των μοριακών πληθυσμών στη αρχή κάθε νέας στοχαστικής επανάληψης.



Εικόνα 6: Φάσεις λειτουργίας FRM πυρήνα και ροή δεδομένων κατά τη διάρκεια ενός κύκλου αντιδράσεων.

Το πρώτο στάδιο της φάσης (1) ελέγχεται από την μονάδα HS η οποία για το σκοπό αυτό χρησιμοποιεί τα σήματα ελέγχου: *Init STI*, *Init RT*, *Init RNG* και *Set Sim Par*. Τα δύο πρώτα σήματα χρησιμοποιούνται για την αρχικοποίηση των πινάκων RT και STI. Κάθε PE μονάδα χρειάζεται μια γεννήτρια ψευδοτυχαίων αριθμών (RNG) για τον υπολογισμό του χρόνου t_{μ} σύμφωνα με τον αλγόριθμο FRM-SSA που περιγράφηκε στην Ενότητα 2.4. Το σήμα *Init RNG* χρησιμοποιείται για την αρχικοποίηση RNG μονάδων της μονάδας FMU. Το σήμα *Set Sim Par* χρησιμοποιείται για την αποθήκευση των παραμέτρων της προσομοίωσης. Το δεύτερο στάδιο ελέγχεται από τη μονάδα SCU μέσω του σήματος *Init Sim* που σηματοδοτεί την αρχικοποίηση των ST πινάκων των FPU μονάδων. Το σήμα *Init Sim* εκκινεί την μονάδα STI η οποία με τη σειρά της αρχικοποιεί τις FPU μονάδες μέσω του σήματος *Init ST*. Τα σήματα *Init ST*, *Init RT*, *Init RNG* οδηγούνται σε μια λογική πύλη OR για να παραχθεί το σήμα *Init* το οποίο στη συνέχεια τροφοδοτείται σε όλες τις FPU μονάδες της FMU μονάδας. Οι FPU μονάδες ελέγχοντας το σήμα *Init* είναι σε θέση να αρχικοποιήσουν τις μνήμης τους με τα δεδομένα που τους αντιστοιχούν.

Η δεύτερη φάση (*Computation* (t_{μ}) – *Phase 2*) της λειτουργίας του FRM πυρήνα περιλαμβάνει τον υπολογισμό του χρόνου t_{μ} και την εύρεση της "νικητήριας αντίδρασης" R_{μ} από τη μονάδα FMU. Η τρίτη φάση (*Update Sim State* – *Phase 3*) περιλαμβάνει την ανανέωση, από την μονάδα STU, των πληθυσμών των μοριακών ειδών σύμφωνα με τη στοιχειομετρία της αντίδρασης R_{μ} και την ανανέωση των χρόνων προσομοίωσης t_{sim} και t_{sam} .

Η φάση (2) ελέγχεται από την μονάδα SCU μέσω του σήματος *Start Sim*. Η μονάδα SCU ξεκινάει μια νέα φάση υπολογισμού όσο ο χρόνος προσομοίωσης t_{sim} είναι μικρότερος του προκαθορισμένου από τον χρήστη χρόνου

T_{sim} . Η φάση (2) ολοκληρώνεται μόλις υπολογιστεί ο χρόνος τ_{μ} από την μονάδα FMU. Με την ολοκλήρωση της φάσης (2) ξεκινάει η τρίτη φάση (*Update Sim State – Phase 3*) η οποία σηματοδοτείται μέσω του σήματος *FMU Valid Out*.

Η φάση (3) ελέγχεται από την μονάδα STU και χωρίζεται σε δύο στάδια. Το πρώτο στάδιο περιλαμβάνει τον υπολογισμό των νέων πληθυσμών (X_{1-8}) των μοριακών ειδών, την ανανέωση των χρόνων t_{sim} και t_{sam} και την ανανέωση του 64-bit μετρητή (Reaction Counter - RC) που χρησιμοποιείται για την καταγραφή των κύκλων αντιδράσεων. Το δεύτερο στάδιο της φάσης (3) περιλαμβάνει την ανανέωση των πληθυσμών (X_{1-8}) των μοριακών ειδών των πινάκων ST του FRM πυρήνα με τους πληθυσμούς που υπολογίστηκαν από την μονάδα STU κατά το πρώτο στάδιο. Το στάδιο αυτό σηματοδοτείται μέσω του σήματος *Update* το οποίο ενημερώνει όλες τις μονάδες του FRM πυρήνα που περιέχουν πίνακες ST ότι θα πρέπει να ανανεώσουν τους πληθυσμούς των πινάκων τους σύμφωνα με τα δεδομένα που παράγονται από την μονάδα STU. Η λήξη της φάσης (3) σηματοδοτείται από την κατερχόμενη παρυφή του σήματος *Update* η οποία ελέγχεται από την μονάδα SCU όπου και αποφασίζεται το πότε θα ξεκινήσει ένας νέος κύκλος αντιδράσεων. Προκειμένου να ανανεωθούν οι πίνακες ST του FRM πυρήνα και οι χρόνοι προσομοίωσης η μονάδα STU μεταδίδει την ακολουθία δεδομένων που παρουσιάζεται στον παρακάτω πίνακα:

Πίνακας 3: Η ακολουθία δεδομένων που παράγεται από την μονάδα STU

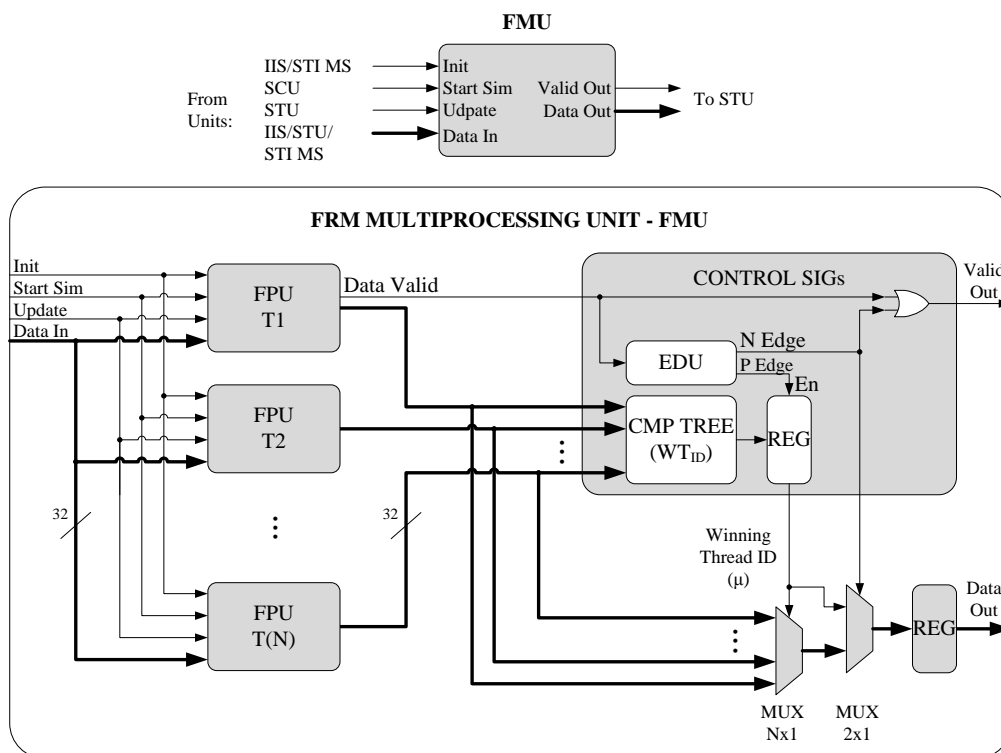
Η ακολουθία δεδομένων που παράγεται από την μονάδα STU – 22 Λέξεις των 32 bit (STU Data Sequence – STU DS)	
SP₁₋₈	Οι 8 δείκτες των μοριακών ειδών (Species Pointers) που συμμετέχουν στην νικητήρια αντίδραση R_{μ} . Χρησιμοποιούνται για την διευθυνσιοδότηση των ανανεωμένων μοριακών ειδών από όλες τις μονάδες του FRM πυρήνα που περιέχουν πίνακες ST.
μ, WT_{ID}	Ο δείκτης της νικητήριας αντίδρασης (μ) και ο δείκτης του νικητήριου νήματος WT_{ID} από την οποία προέκυψε. Οι δείκτες αυτοί μεταδίδονται από την μονάδα OIS στο Host PC.
X_{1-8}	Οι 8 ανανεωμένοι πληθυσμοί (X_i) των μοριακών ειδών. Χρησιμοποιούνται για την ανανέωση των πληθυσμών των μοριακών ειδών από όλες τις μονάδες του FRM πυρήνα που περιέχουν πίνακες ST.
t_{sim}, t_{sam}	Οι χρόνοι t_{sim} και t_{sam} χρησιμοποιούνται από την μονάδα SCU για τον έλεγχο της προσομοίωσης ενώ μεταδίδονται και στο Host PC μέσω της μονάδας OIS.
RC_{H-L}	Δύο λέξεις για το υψηλό (high) και το χαμηλό (low) κομμάτι του 64-bit μετρητή RC. Μεταδίδονται στο Host PC από την μονάδα OIS.

Παράλληλα με τις παραπάνω φάσεις πραγματοποιείται και η μετάδοση των αποτελεσμάτων της προσομοίωσης μέσω της μονάδας OIS. Η μονάδα OIS μεταδίδει τα αποτελέσματα της προσομοίωσης είτε εφαρμόζοντας

δειγματοληψίας σε αυτά είτε όχι, ανάλογα με την προκαθορισμένη από την χρήστη παράμετρο T_{sam} . Κατά την μετάδοση των αποτελεσμάτων με δειγματοληψία, στο Host PC στέλνονται οι πληθυσμοί των μοριακών ειδών που μεταβλήθηκαν στην τελευταία περίοδο δειγματοληψίας. Επίσης στέλνεται και η τιμή της προηγούμενης περιόδου δειγματοληψίας. Η διαδικασία αυτή σηματοδοτείται από την μονάδα SCU μέσω του σήματος *New SP* (New Sampling Period). Κατά την μετάδοση χωρίς δειγματοληψίας, στο Host PC στέλνονται οι ακόλουθες τιμές του παραπάνω πίνακα: ο δείκτης (μ) της νικητήριας αντίδρασης R_μ , ο δείκτης του νικητήριου νήματος WT_{ID} , οι 8 ανανεωμένοι πληθυσμοί (X_i) των μοριακών ειδών και η νέος χρόνος προσομοίωσης t_{sim} . Και στις δύο περιπτώσεις, με το τέλος μιας στοχαστικής επανάληψης μεταδίδεται στο Host PC μέσω της μονάδας OIS η τιμή του RC μετρητή. Η αποστολή της τιμής του μετρητή σηματοδοτείται από την μονάδα SCU μέσω του σήματος *Send RC* (Send Reaction Counter).

4.4.1 FRM Multiprocessing Unit - FMU

Η μονάδα FMU είναι υπεύθυνη για τη παράλληλη επεξεργασία ενός βιολογικού μοντέλου. Η παραλληλοποίηση του αλγορίθμου επιτυγχάνεται μέσω των επεξεργαστικών μονάδων (FRM Processing Units - FPU) που αναλαμβάνουν να επεξεργαστούν παράλληλα, υποσύνολα αντιδράσεων από το σύνολο των αντιδράσεων ενός μοντέλου. Ο αριθμός των FPUs καθορίζεται από τη Generic παράμετρο N η οποία ορίζεται στο top-level αρχείο της σχεδίασης του FRM SoC. Για λόγους παραμετροποίησης σε επίπεδο υλοποίησης η παράμετρος N μπορεί να λάβει μόνο τιμές που είναι δυνάμεις του δύο ενώ η μέγιστη τιμή περιορίζεται μόνο από τους διαθέσιμους πόρους του FPGA. Στη συνέχεια ακολουθεί σχηματικό αρχιτεκτονικής της μονάδας FMU.



Εικόνα 7: Αρχιτεκτονική της μονάδας FRM Multiprocessing Unit - FMU

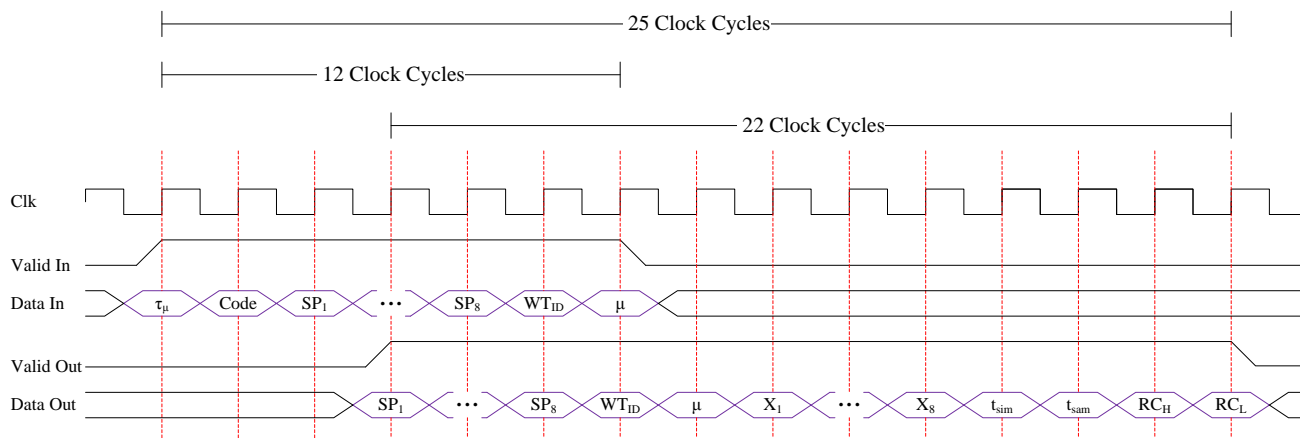
Κάθε FPU αποτελεί ένα νήμα υλικού (thread) που αναλαμβάνει το υπολογισμό ενός τμήματος του πίνακα RT. Με το τέλος ενός κύκλου αντιδράσεων, κάθε FPU παράγει ένα χρόνο τ_μ οποίος όμως θα έχει προκύψει από το υποσύνολο των αντιδράσεων που επεξεργάστηκε η FPU και όχι από το πλήρες σύνολο των αντιδράσεων του μοντέλου. Προκειμένου να εντοπιστεί ο συνολικά μικρότερος χρόνος τ_μ σχεδιάστηκε και υλοποιήθηκε ένα δέντρο συγκριτών (CMP TREE) το οποίο είναι επίσης παραμετροποιήσιμο σύμφωνα με το παράμετρο N . Στις N εισόδους του δέντρου θα οδηγηθούν όλοι οι χρόνοι τ_μ που υπολογίστηκαν από τις διαθέσιμες FPU. Έπειτα, το δέντρο συγκριτών μετά από $\log_2(N)$ κύκλους ρολογιού θα τροφοδοτήσει στην έξοδο του το αναγνωριστικό του νήματος που συνεισέφερε συνολικά μικρότερο χρόνο τ_μ . Το αναγνωριστικό αυτό, στην Εικόνα 4.3 αναφέρεται και ως Winning Thread ID - WT_{ID}. Στη συνέχεια, αφού πρώτα αποθηκευθεί σε ένα καταχωρητή, το αναγνωριστικό του νήματος χρησιμοποιείται ως σήμα επιλογής του πολυπλέκτη MUX $N \times 1$. Με αυτό τον τρόπο επιτυγχάνεται η πολυπλεξία των FPUs μονάδων αφού την έξοδο της FMU την οδηγεί μόνο η FPU που υπολόγισε το συνολικά μικρότερο χρόνο τ_μ .

Όπως αναφέρθηκε η τιμή Winning Thread ID αποθηκεύεται σε έναν καταχωρητή. Προκειμένου ο καταχωρητής να ενεργοποιηθεί τη σωστή χρονική στιγμή χρησιμοποιείται το σήμα Data Valid της πρώτης FPU μονάδας. Το σήμα Data Valid των FPUs παίρνει τη λογική τιμή (1) μόλις η μονάδα COMP TREE εντοπίσει το νικητήριο νήμα και παραμένει σε αυτή τη κατάσταση για όσους κύκλους ρολογιού η μονάδα FPU παράγει έγκυρα δεδομένα. Η ακολουθία των δεδομένων που παράγονται από μια FPU, μετά το τέλος ενός κύκλου αντιδράσεων, σχετίζεται με το χρόνο τ_μ που υπολόγισε, τη στοιχειομετρία της αντίδρασης R_μ , το δείκτη προς την αντίδραση R_μ και δείκτες προς τα αντιδρώντα και τα προϊόντα της αντίδρασης R_μ . Το σήμα Data Valid της πρώτης FPU μονάδας οδηγείται σε μια μονάδα ανίχνευσης ακμών σήματος (Edge Detector Unit - EDU). Από τη μονάδα EDU παράγονται δύο παλμοί, ένας παλμός κατά την ανερχόμενη παρυφή του σήματος Data Valid και ένας κατά την κατερχόμενη παρυφή. Ο παλμός που παράγεται κατά την ανερχόμενη παρυφή του σήματος Data Valid χρησιμοποιείται ως το σήμα ενεργοποίησης (Enable) του καταχωρητή που αποθηκεύει το Winning Thread ID. Ο παλμός που παράγεται κατά την κατερχόμενη παρυφή λειτουργεί ως το σήμα επιλογής ενός πολυπλέκτη δύο σε ένα (MUX 2×1). Ο πολυπλέκτης αυτός είναι απαραίτητος προκειμένου να οδηγηθεί στην έξοδο της FMU και το αναγνωριστικό του νικητήριου νήματος καθώς χρησιμοποιείται και από άλλες μονάδες του FRM πυρήνα.

Ο λόγος που χρησιμοποιείται το σήμα Data Valid μόνο της πρώτης FPU είναι ότι όλες οι FPU είναι προγραμματισμένες να ολοκληρώνουν τους υπολογισμούς του ταυτόχρονα. Κάτι τέτοιο επιτυγχάνεται εισάγοντας επιπλέον αντιδράσεις, που δεν επηρεάζουν σε καμία περίπτωση το πείραμα, στο πίνακα RT προκειμένου το βάθος του να είναι πάντοτε δύναμη του δύο ανεξάρτητα από το αρχικό πλήθος των αντιδράσεων ενός βιομοντέλου.

4.4.2 Species/Time Update Unit - STU

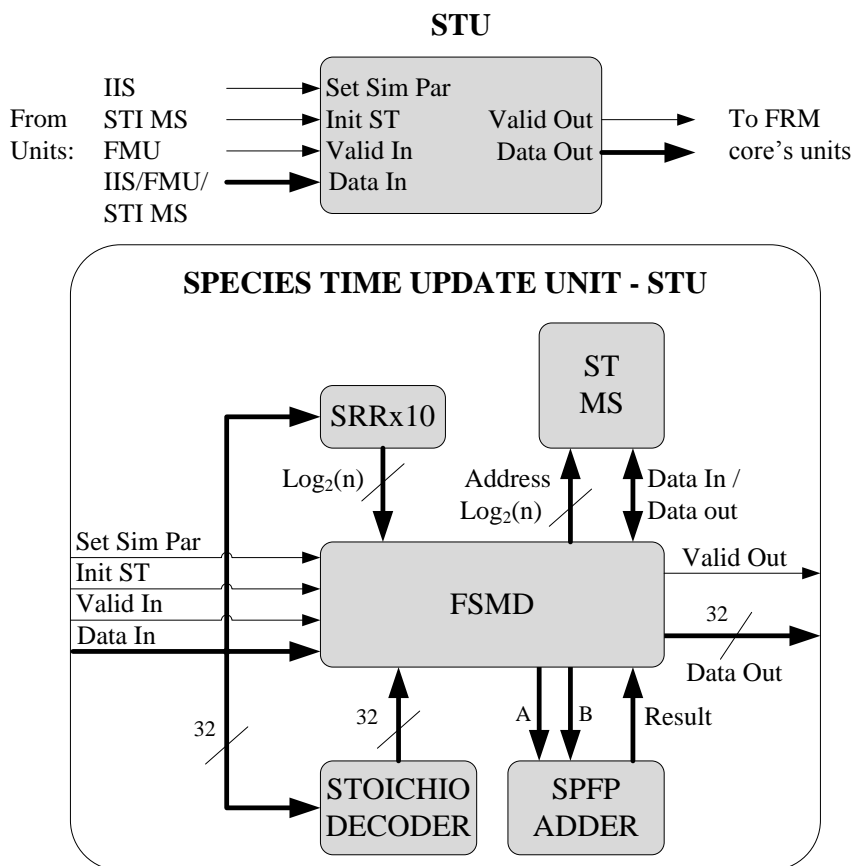
Η μονάδα STU είναι υπεύθυνη για την ανανέωση των μοριακών πληθυσμών της R_μ νικητήριας αντίδρασης, του χρόνου προσομοίωσης και του χρόνου δειγματοληψίας. Η λειτουργία και η συμπεριφορά της μονάδας STU μπορεί να συνοψιστεί στο παρακάτω διάγραμμα κυματομορφών.



Εικόνα 8: Διάγραμμα κυματομορφών της μονάδας SPU

Όπως φαίνεται και στη παραπάνω εικόνα, η μονάδα STU δέχεται ως είσοδο την ακολουθία δεδομένων που παράγεται από τη μονάδα FMU. Η ακολουθία αυτή περιλαμβάνει δεδομένα όπως: ο χρόνος τ_{μ} , η κωδικοποίηση της στοιχειομετρίας (Code) της αντίδρασης R_{μ} , οι οχτώ δείκτες προς μοριακά είδη (SP_i) της αντίδρασης R_{μ} , ένα δείκτη προς την αντίδραση R_{μ} και το αναγνωριστικό του νικητήριου νήματος (μ). Στην έξοδο της μονάδας STU τροφοδοτούνται δεδομένα όπως: οι οχτώ δείκτες προς μοριακά είδη της αντίδρασης R_{μ} προκειμένου να χρησιμοποιηθούν και από άλλες μονάδες του FRM πυρήνα, ο δείκτης προς την αντίδραση R_{μ} , το αναγνωριστικό του νικητήριου νήματος, οι οχτώ ανανεωμένοι πληθυσμοί μοριακών ειδών (X_i), ο νέος χρόνος προσομοίωσης (t_{sim}), ο νέος χρόνος δειγματοληψίας (t_{sam}) και η 64-bit τιμή ενός μετρητή (RC_{HIGH} και RC_{LOW}) που καταγράφει το συνολικό αριθμό κύκλων αντιδράσεων. Η λειτουργία της μονάδας STU και της φάσης ανανέωσης του συστήματος ολοκληρώνεται σε 25 κύκλους ρολογιού.

Για να επιτευχθούν οι προαναφερθείσες λειτουργίες η μονάδα STU χρησιμοποιεί μια μνήμη για την αποθήκευση των μοριακών πληθυσμών, έναν Single Precision Floating Point - SPFP αθροιστή έναν αποκωδικοποιητή στοιχειομετρίας, καταχωρητές ολίσθησης και μια FSMD για το συντονισμό όλων των παραπάνω υπομονάδων. Στη συνέχεια ακολουθεί σχηματικό αρχιτεκτονικής της μονάδας STU.



Εικόνα 9: Αρχιτεκτονική της μονάδας STU

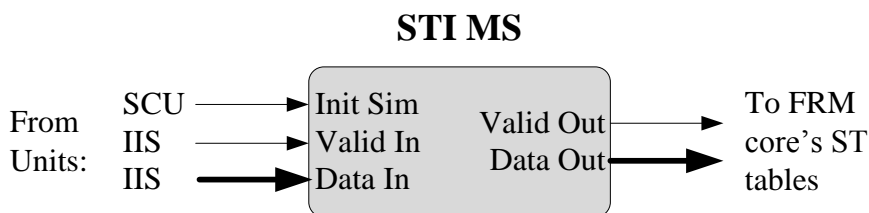
Ο SPFP ADDER αναλαμβάνει να εκτελέσει όλες τις Floating Point αθροίσεις για να υπολογιστούν οι νέοι μοριακοί πληθυσμοί και οι χρόνοι προσομοίωσης. Η μνήμη ST έχει αποθηκευμένους όλους τους μοριακούς πληθυσμούς και χρησιμοποιεί τους δείκτες SP_i για την εύρεση των μοριακών πληθυσμών της νικητώρας αντίδρασης R_μ . Η μνήμη ST ανανεώνεται και αυτή με τις νέες τιμές πληθυσμών που θα υπολογιστούν από τον SPFP ADDER. Ο SRR μήκους 10 θέσεων χρησιμοποιείται προκειμένου τη κατάλληλη χρονική στιγμή οι δείκτες SP_i να αρχίσουν να τροφοδοτούνται στην έξοδο της μονάδας STU. Σχετικά με τη μονάδα αποκωδικοποίησης της στοιχειομετρίας (STOICHIO DECODER), δέχεται ως είσοδο μια λέξη των 32-bit (Code) όπου σε αυτή έχουν κωδικοποιηθεί οι στοιχειομετρίες των οχτώ μοριακών ειδών μιας αντίδρασης και σε οχτώ κύκλους ρολογιού παράγει 8 λέξεις των 32-bit που αντιστοιχούν στη στοιχειομετρία της αντίδρασης.

Τέλος θα πρέπει να αναφερθεί ότι η σχεδίαση της μονάδας STU έγινε πολύ προσεκτικά παίρνοντας υπόψη παραμέτρους όπως: ο λανθάνων χρόνος (latency) του SPFP αθροιστή, ο συνολικός λανθάνων χρόνος της μονάδας STU, ο αριθμός των λέξεων που μπορούν να εγγραφούν ταυτόχρονα στους πίνακες ST του FRM πυρήνα καθώς και παραμέτρους που σχετίζονται με τις απαιτήσεις σε πόρους υλικού της μονάδας STU.

4.4.3 ST Initialization Memory System – STI MS

Στην αρχή μιας επανάληψης (repetition) προσομοίωσης οι πίνακες ST του FRM πυρήνα πρέπει να αρχικοποιηθούν με τους αρχικούς πληθυσμούς των μοριακών ειδών. Η μονάδα STI MS (ST Initialization

Memory System) είναι υπεύθυνη για αυτήν την αρχικοποίηση. Στη συνέχεια ακολουθεί σχηματικό μπλοκ διαγράμματος με τα σήματα εισόδου και εξόδου της μονάδας STI MS.



Εικόνα 10: Μπλοκ διάγραμμα της μονάδας STI MS

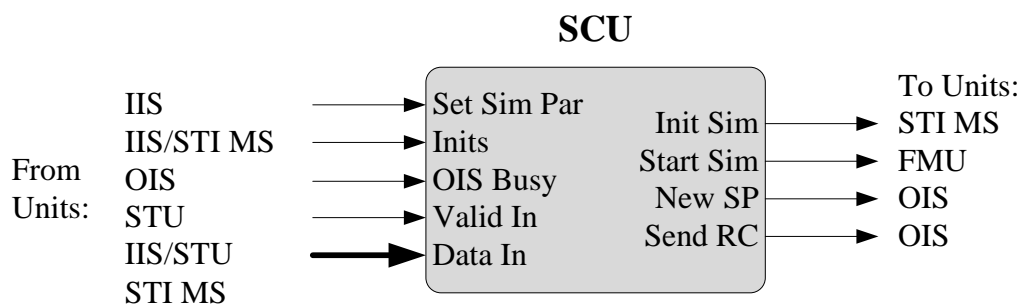
Το σήμα *Valid In* οδηγείται από τη μονάδα IIS. Μέσω του σήματος *Valid In* ενημερώνεται η μονάδα STI MS ότι καταφθάνουν έγκυρα δεδομένα (οι αρχικοί πληθυσμοί των μοριακών ειδών) στην είσοδο δεδομένων της (*Data In*) από το δίκτυο. Το σήμα *Init Sim* οδηγείται από τη μονάδα SCU και σηματοδοτεί την έναρξη της αρχικοποίησης. Κατά το στάδιο αυτό η μονάδα STI MS μεταδίδει μέσω του σήματος *Data Out* τους αρχικούς πληθυσμούς των μοριακών ειδών στις μονάδες του FRM πυρήνα που διαθέτουν πίνακες ST. Οι μονάδες αυτές ενημερώνονται για την έναρξη της αρχικοποίησης μέσω του σήματος *Valid Out*.

Η μονάδα STI MS αποτελείται από μια μνήμη τυχαίας προσπέλασης με μια θύρα για γράψιμο και διάβασμα δεδομένων (Single Port Random Access Memory – SPRAM) βάθους N λέξεων των 32-bit. Στην SPRAM της μονάδας STU κατά το στάδιο της αρχικοποίησης του FRM πυρήνα αποθηκεύονται οι αρχικοί πληθυσμοί των μοριακών ειδών ενός μοντέλου. Έπειτα, η μνήμη SPRAM της μονάδας STI MS χρησιμοποιείται για την αρχικοποίηση όλων των πινάκων ST του FRM πυρήνα. Κατά τη διάρκεια της εκτέλεσης μιας προσομοίωσης η μνήμη SPRAM δεν ανανεώνεται και έτσι δίνεται η δυνατότητα αρχικοποίησης των πινάκων ST του πυρήνα μετά το τέλος μιας προσομοίωσης. Ουσιαστικά η μνήμη SPRAM της μονάδας STI MS δρα ως μια κρυφή μνήμη (Cache Memory) και προσφέρει τη δυνατότητα εκτέλεσης πολλών στοχαστικών επαναλήψεων μιας προσομοίωσης χωρίς να απαιτείται η επικοινωνία του FRM SoC με το Host PC για την επαναρχικοποίηση των πινάκων ST.

Για την διευθυνσιοδότηση της μνήμης SPRAM και κατά την εγγραφή αλλά και κατά το διάβασμα χρησιμοποιούνται εσωτερικοί μετρητές. Μέσω του σήματος *Valid In* ενεργοποιείται ο μετρητής εγγραφών και μέσω του σήματος *Init Sim* ενεργοποιείται ο μετρητής ανάγνωσης των μοριακών πληθυσμών. Όλα τα συστήματα μνήμης (memory systems) του FRM πυρήνα διαθέτουν αντίστοιχους μετρητές εγγραφών και ανάγνωσης. Η σχεδιαστική επιλογή αυτή έγινε προκειμένου οι διεπαφές των μονάδων να έχουν όσο το δυνατόν λιγότερους διαύλους δεδομένων ενώ παράλληλα με αυτόν τον τρόπο κρατείται στο ελάχιστο το μέγεθος "ικανότητα οδήγησης πύλης" (fan-out) του συστήματος.

4.4.4 Simulation Control Unit – SCU

Η μονάδα SCU είναι υπεύθυνη για τον έλεγχο και την εκτέλεση προσομοιώσεων στον FRM πυρήνα. Για τη VHDL υλοποίηση της μονάδας σχεδιάστηκε και αναπτύχθηκε μια μηχανή FSM. Στη συνέχεια ακολουθεί σχηματικό μπλοκ διαγράμματος με τα σήματα εισόδου και εξόδου της μονάδας STI MS.



Εικόνα 11: Μπλοκ διάγραμμα SCU μονάδας

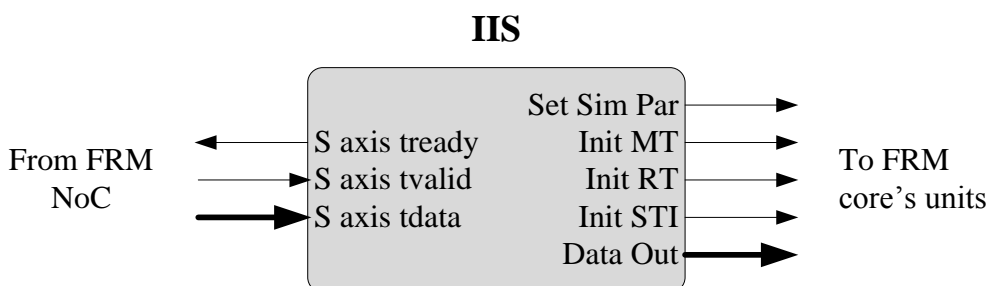
Οι αρμοδιότητες της μονάδας SCU συνοψίζονται από τα σήματα εξόδου της. Το σήμα *Init Sim* (initialize simulation) τροφοδοτείται στη μονάδα STI MS και μέσω ενός παλμού σηματοδοτεί την αρχικοποίηση του συστήματος στην αρχή μίας νέας επανάληψης της προσομοίωσης. Το σήμα *Start Sim* (start simulation) σηματοδοτεί την έναρξη μιας νέας επανάληψης της προσομοίωσης και τροφοδοτείται στη μονάδα FMU. Η οδήγηση των παραπάνω σημάτων εξαρτάται από τρεις βασικούς ελέγχους που πραγματοποιεί η μονάδα SCU. Ο πρώτος έλεγχος σχετίζεται με τον αν ο χρόνος t_{sim} έχει φτάσει το μέγιστο χρόνο προσομοίωσης T_{sim} μιας επανάληψης. Ο δεύτερος έλεγχος σχετίζεται με το αν έχουν ολοκληρωθεί όλες οι επαναλήψεις (R) των προσομοιώσεων που έχει ορίσει ο χρήστης. Ο τρίτος έλεγχος σχετίζεται με το αν το σύστημα μπορεί να προχωρήσει στην έναρξη ενός νέου κύκλου αντιδράσεων (RC) λόγω της μονάδας Output Interface System – OIS. Κάτι τέτοιο μπορεί να συμβεί στη περίπτωση που η μονάδα OIS είναι απασχολημένη με τη μετάδοση δεδομένων που έχουν προκύψει από προηγούμενους κύκλων αντιδράσεων και το σύστημα δεν είναι σε θέση να αποθηκεύσει τα νέα δεδομένα που θα παραχθούν.

Το σήμα *New SP* (new sampling period) σηματοδοτεί το τέλος μιας περιόδου δειγματοληψίας και τροφοδοτείται στη μονάδα OIS για να την ενημερώσει ότι θα πρέπει να στείλει στο Host PC τα δεδομένα του πίνακα ST που έχουν μεταβληθεί κατά τη τελευταία περίοδο δειγματοληψίας. Κάτι τέτοιο θα συμβεί όταν ο χρόνος προσομοίωσης t ξεπεράσει το χρόνο δειγματοληψίας t_{sam} . Στη περίπτωση αυτή στέλνονται τα δεδομένα του πίνακα ST στο Host PC και ανανεώνεται και ο χρόνος δειγματοληψίας $t_{sampling}$ ($t_{sampling} = t_{sam} + T_{sam}$). Το σήμα *Send RC* (send reaction counter) σηματοδοτεί το τέλος μιας επανάληψης της προσομοίωσης και τροφοδοτείται στη μονάδα OIS για να την ενημερώσει ότι θα πρέπει να σταλεί στο Host PC η τιμή του 64-bit μετρητή που καταγράφει τους κύκλους αντιδράσεων που χρειάστηκαν προκειμένου να ολοκληρωθεί η προσομοίωση. Όπως ήδη έχει αναφερθεί, υπεύθυνη για την αύξηση αυτού του μετρητή είναι η FSMD της μονάδας STU.

Στην Ενότητα 2 αναφέρθηκε ότι αν οι πληθυσμοί των μοριακών ειδών αποκτήσουν υψηλές τιμές τότε τα χρονικά βήματα (τ_μ) που παράγονται από τους αλγόριθμους του Gillespie μπορούν να προκύψουν αρκετά μικρά. Το φαινόμενο αυτό διαχειρίζεται η SCU μονάδα καταγράφοντας μέσω ενός 32-bit μετρητή τους κύκλους αντιδράσεων ανάμεσα σε δύο περιόδους δειγματοληψίας. Στη περίπτωση όπου αυτός ο μετρητής ξεπεράσει μια τιμή προκαθορισμένη από τον χρήστη (Max Internal Steps) τότε η συγκεκριμένη επανάληψη τερματίζεται και το σύστημα προχωρά στην επόμενη επανάληψη προσομοίωσης, εάν υπάρχει. Τον συγκεκριμένο τρόπο αντιμετώπισης τον έχουμε δανειστεί από το δημοφιλή προσομοιωτή βιολογικών δικτύων COPASI [53].

4.4.5 Input Interface System - IIS

Στον Πίνακα 1 δόθηκαν τα σήματα της διεπαφής του FRM SoC. Όπως είναι αναμενόμενο, λόγω του πρωτοκόλλου του δικτύου, τα ίδια σήματα αποτελούν και τη διεπαφή ενός FRM πυρήνα. Για την οδήγηση αυτών των σημάτων σχεδιάστηκαν και υλοποιήθηκαν δύο ξεχωριστές μονάδες, IIS και OIS, όπου η μονάδα IIS ασχολείται με τη μεταφορά δεδομένων προς τον FRM πυρήνα ενώ η μονάδα OIS ασχολείται με την αποστολή δεδομένων προς το δίκτυο. Για τη VHDL υλοποίηση της μονάδας IIS σχεδιάστηκε και αναπτύχθηκε μια μηχανή FSMD η οποία έχει δύο βασικές λειτουργίες. Η πρώτη λειτουργία σχετίζεται με το να ζητήσει από το δίκτυο όλα τα απαραίτητα δεδομένα που απαιτούνται προκειμένου να ξεκινήσει μια προσομοίωση. Η επικοινωνία αυτή στηρίζεται, όπως έχει ήδη αναφερθεί, στο πρωτόκολλο AXIS4 συνεπώς είναι υπεύθυνη για τον έλεγχο των σημάτων: *s axis tvalid*, *s axis tid* και *s axis tready*. Η δεύτερη λειτουργία σχετίζεται με το να ενημερώνει τις κατάλληλης χρονικές στιγμές τις υπόλοιπες μονάδες του FRM πυρήνα για πότε πρέπει να αποθηκεύσουν τα δεδομένα που καταφθάνουν στην είσοδο του FRM πυρήνα. Για το λόγο αυτό χρησιμοποιεί τα σήματα: *set sim setup*, *init mt*, *init rt* και *init sti*. Στη συνέχεια ακολουθεί σχηματικό μπλοκ διαγράμματος με τα σήματα εισόδου και εξόδου της μονάδας IIS.



Εικόνα 12: Μπλοκ διάγραμμα της μονάδας IIS

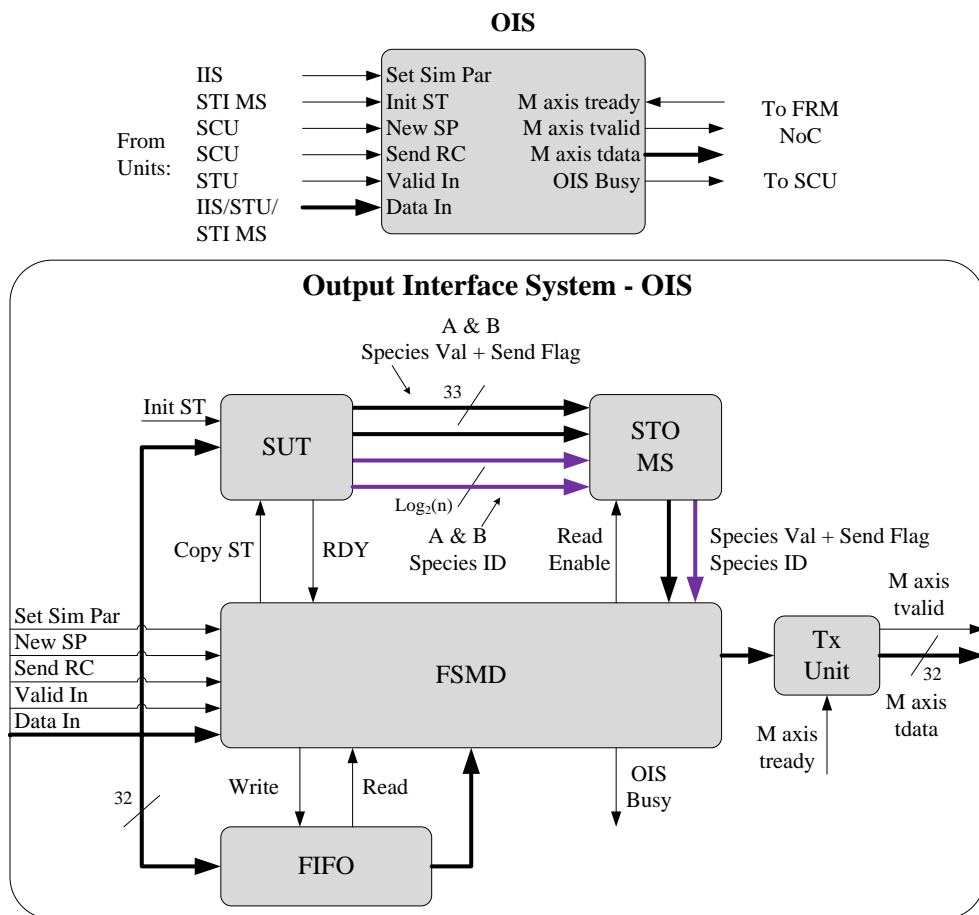
Τα σήματα *S axis tvalid*, *S axis tready* και *S axis tdata* σχετίζονται απόλυτα με το πρωτόκολλο AXIS4. Τα σήματα *Set Sim Par*, *Init MT*, *Init RT* και *Init STI* παίρνουν λογικές τιμές (1) για όσο μεταφέρονται δεδομένα από το δίκτυο προς τον πυρήνα και σχετίζονται είτε με τις παραμέτρους είτε με τα δεδομένα της προσομοίωσης. Ελέγχοντας τα σήματα αυτά οι μονάδες του FRM πυρήνα είναι σε θέση να αποθηκεύσουν τα δεδομένα και τις παραμέτρους που χρειάζονται για την εκτέλεση της προσομοίωσης.

Σχετικά με τη λειτουργία της μονάδας IIS, η FSMD της μονάδας αρχικά περιμένει να λάβει όλη τη πληροφορία που βρίσκεται στο αρχείο των παραμέτρων (PAR) του οποίου η δομή του ορίστηκε στην Ενότητα 3.3. Το αρχείο των παραμέτρων είναι σταθερού μεγέθους και έτσι η FSMD γνωρίζει εκ των προτέρων πόσες τιμές πρέπει να αναμένει στη φάση αυτή. Μόλις λάβει το αρχείο των παραμέτρων, η διαδικασία αποστολής δεδομένων συνεχίζεται με την αποστολή του πίνακα αρχικοποίησης των μηχανών παραγωγής ψευδοτυχαίων αριθμών. Και οι πίνακες αυτοί είναι σταθερού μεγέθους. Τέλος, αναμένεται η αποστολή του πίνακα RT και ST. Οι πίνακες αυτοί δεν είναι σταθερού μεγέθους αλλά εξαρτούνται από το εκάστοτε μοντέλο προς προσομοίωση. Ωστόσο, εφόσον έχει ήδη ληφθεί το αρχείο των παραμέτρων η FSMD είναι σε θέση πλέον να γνωρίζει το μέγεθος των πινάκων RT και ST και έτσι ολοκληρώνεται η διαδικασία αποστολής δεδομένων στον FRM πυρήνα.

4.4.6 Output Interface System - OIS

Η μονάδα OIS είναι υπεύθυνη για την αποστολή δεδομένων προς το δίκτυο και κατ' επέκταση προς το Host PC. Συνεπώς, η μονάδα OIS είναι εκείνη που ελέγχει τα σήματα τύπου Master της διεπαφής του FRM πυρήνα. Τα σήματα αυτά είναι τα ακόλουθα: *M axis tready*, *M axis tvalid* και *M axis tdata*. Η λειτουργία της μονάδας OIS καθορίζεται από τα σήματα εισόδου *New SP* και *Send RC* τα οποία οδηγούνται από τη μονάδα SCU.

Η αποστολή δεδομένων υποστηρίζεται με δύο διαφορετικούς τρόπους (Modes). Κατά τον ένα τρόπο, στέλνεται η ακολουθία των δεδομένων που παράγεται από τη μονάδα SPU, χωρίς όμως τους δείκτες (SP_i), μετά το τέλος κάθε κύκλου αντιδράσεων. Ο λόγος που δεν στέλνονται οι δείκτες SP_i είναι ότι μέσω του δείκτη της R_μ αντίδρασης είναι εφικτή η γραφική αναπαράσταση των αποτελεσμάτων. Ωστόσο, όπως αναφέρθηκε και στην Ενότητα 3, κάτι τέτοιο οδηγεί σε τεράστιο όγκο δεδομένων ο οποίος θα πρέπει να σταλθεί και να αποθηκευθεί στο Host PC. Για το λόγο αυτό καταλήξαμε στη λύση της δειγματοληψίας των αποτελεσμάτων που αποτελεί και τον δεύτερο και κυρίαρχο τρόπο λειτουργίας. Κατά αυτόν τον τρόπο η αποστολή των δεδομένων γίνεται μόνο όταν ολοκληρωθεί μια περίοδος δειγματοληψίας και αφορά μόνο τους πληθυσμούς των μοριακών ειδών που έχουν μεταβληθεί κατά τη διάρκεια της. Στη συνέχεια ακολουθεί σχηματικό αρχιτεκτονικής της μονάδας OIS.



Εικόνα 13: Αρχιτεκτονική της μονάδας OIS

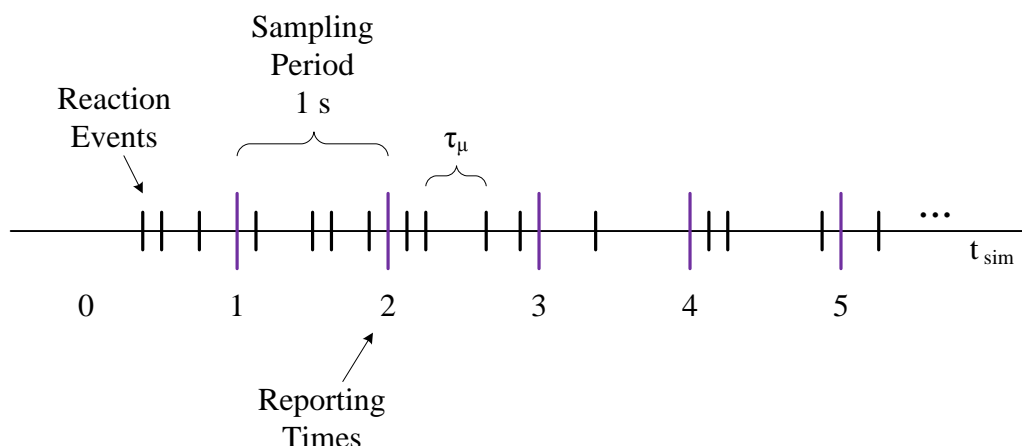
Η μονάδα OIS αποτελείται από τέσσερις υπομονάδες που είναι υπεύθυνες για τη διεκπεραίωση συγκεκριμένων λειτουργιών και μια FSMD για το συντονισμό τους. Μετά το τέλος κάθε κύκλου αντιδράσεων πραγματοποιείται η ανανέωση των μοριακών πληθυσμών. Οι ανανεωμένοι πληθυσμοί καταγράφονται στη υπομονάδα Species Update Table –SUT. Επιπλέον σε αυτή καταγράφονται και το ποιοί μοριακοί πληθυσμοί έχουν ανανεωθεί. Για αυτό το λόγο, η λέξη του πίνακα ST που βρίσκεται στη μονάδα SUT έχει εύρος 33-bit. Το πιο σημαντικό ψηφίο (MBS) κάθε λέξης χρησιμοποιείται ως σημαία που σηματοδοτεί ότι η τιμή του συγκεκριμένου είδους έχει αλλάξει τουλάχιστον μια φορά κατά τη τρέχουσα περίοδο δειγματοληψίας.

Με την ολοκλήρωση μιας περιόδου δειγματοληψίας η μονάδα SCU εκκινεί τη διαδικασία αποστολής δεδομένων της μονάδας OIS μέσω του σήματος ελέγχου *New SP*. Το πρώτο στάδιο της διαδικασίας αυτής περιλαμβάνει την αντιγραφή των n στοιχείων του πίνακα ST της μονάδας SUT στο πίνακα ST της μονάδας Species Table Output Memory System - STO MS. Ο λόγος που γίνεται αυτό είναι για να μπορέσει το σύστημα να προχωρήσει στον επόμενο κύκλο αντιδράσεων (reaction cycle) ενώ ταυτόχρονα η μονάδα OIS μεταδίδει τα δεδομένα του προηγούμενου κύκλου αντιδράσεων (latency hiding).

Θα πρέπει να αναφερθεί ότι ο χρόνος που μπορεί να μεσολαβήσει για τη διαδικασία της μετάδοσης εξαρτάται από πολλές παραμέτρους του συστήματος (host PC hard drive, host pc OS, PCIe core κ.α.). Συνεπώς, ο FRM πυρήνας, εάν δε μεσολαβούσε το στάδιο της αντιγραφής, θα έπρεπε να περιμένει να ολοκληρωθεί η διαδικασία αυτή προτού προχωρήσει σε νέο κύκλο αντιδράσεων. Η αναμονή αυτή αποφεύγεται εκτός και αν ολοκληρωθεί και δεύτερη περίοδος δειγματοληψίας ενώ ακόμη στέλνονται τα δεδομένα της δειγματοληψίας της προηγούμενης.

Το στάδιο της αντιγραφής σηματοδοτείται από το σήμα *Copy ST* που δέχεται η μονάδα SUT ενώ όσο διαρκεί αυτό το στάδιο ο FRM πυρήνας δεν μπορεί να προχωρήσει στον επόμενο κύκλο αντιδράσεων. Ο έλεγχος αυτός γίνεται από τη μονάδα ελέγχου SCU η οποία προτού ξεκινήσει έναν νέο κύκλο αντιδράσεων ελέγχει το σήμα *OIS Busy* της μονάδας OIS. Η φάση της αντιγραφής διαρκεί $n/2$ κύκλους ρολογιού όπου n ο αριθμός των ειδών ενός μοντέλου. Κάτι τέτοιο είναι εφικτό εφόσον γίνονται δύο εγγραφές ανά κύκλο ρολογιού στους πίνακες ST των μονάδων SUT και STO MS. Για το λόγο αυτό η υλοποίηση των πινάκων ST έχει γίνει με τη χρήση True Dual Port BRAMs. Ταυτόχρονα με την αντιγραφή του πίνακα ST στη μονάδα STO MS γίνεται και η αρχικοποίηση όλων των πιο σημαντικών ψηφίων (MSBs) που χρησιμοποιούνται ως σημαίες για τη μεταβολή των μοριακών πληθυσμών στη μονάδα SUT προκειμένου να μπορούν να χρησιμοποιηθούν στην επόμενη περίοδο δειγματοληψίας.

Μόλις ολοκληρωθεί η αντιγραφή των μοριακών ειδών η FSMD προχωράει στη φάση του διαβάσματος των δεδομένων της μονάδας STO MS ενώ παράλληλα το σύστημα μπορεί να ξεκινήσει τον επόμενο κύκλο αντιδράσεων. Κατά το στάδιο αυτό τα δεδομένα της μονάδας STO MS διαβάζονται σειριακά όπου και ελέγχεται το πιο σημαντικό ψηφίο (MSB) της κάθε λέξης. Εάν το πιο σημαντικό ψηφίο (MSB) έχει τη λογική τιμή (1) τότε η FSMD μεταδίδει προς το δίκτυο, μέσω της μονάδας Tx Unit, το δείκτη και τη τιμή του μοριακού είδους. Η μονάδα Tx Unit είναι υπεύθυνη για τη μετάδοση λέξεων των 32-bit στο δίκτυο του SoC σύμφωνα με τις προδιαγραφές του AXIS4 πρωτοκόλλου. Η παραπάνω διαδικασία ολοκληρώνεται όταν διαβαστούν όλα τα μοριακά είδη του πίνακα ST της μονάδας STO MS των οποίων οι πληθυσμοί άλλαξαν. Η συμπεριφορά του μηχανισμού δειγματοληψίας σε σχέση με τους κύκλους αντιδράσεων παρουσιάζεται στο παρακάτω διάγραμμα.



Εικόνα 14: Διάγραμμα μηχανισμού δειγματοληψίας

Το παραπάνω σχηματικό δείχνει ένα τυχαίο αλλά αντιπροσωπευτικό παράδειγμα του μηχανισμού δειγματοληψίας σε σχέση με τις αντιδράσεις. Στον οριζόντιο άξονα αναπαρίσταται ο χρόνος προσομοίωσης t_{sim} . Οι κάθετες (μαύρες) γραμμές αναπαριστούν τις αντιδράσεις. Τα χρονικά διαστήματα τ_μ μεταξύ δύο αντιδράσεων δεν είναι σταθερά αλλά εξαρτούνται από την νικητήρια αντίδραση R_μ . Οι μεγαλύτερες γραμμές αναπαριστούν τα χρονικά σημεία όπου τελειώνει μια περίοδος δειγματοληψίας και η καταγραφή των μοριακών πληθυσμών προκειμένου να σταλθούν στο Host PC από τη μονάδα OIS. Στο παράδειγμα της Εικόνα 14 η περίοδος δειγματοληψίας (Sampling Period) έχει οριστεί στο ένα δευτερόλεπτο. Επιπλέον, από το σχηματικό παρατηρούμε ότι μέσα σε μια περίοδο δειγματοληψίας μπορεί να συμβούν μια ή και πολλές αντιδράσεις. Σε κάθε περίπτωση όμως θα σταλθούν μόνο εκείνοι οι μοριακοί πληθυσμοί που μεταβλήθηκαν κατά τη τελευταία περίοδο και μόνο η πλέον πρόσφατη (τελευταία) ανανέωση τους. Με αυτό τον τρόπο επιτυγχάνεται ο περιορισμός του παραγόμενου όγκου δεδομένων ενώ παράλληλα επιλέγοντας προσεκτικά τη περίοδο δειγματοληψίας ο χρήστης μπορεί να έχει μια πολύ λεπτομερή εικόνα της εξέλιξης της δυναμικής συμπεριφοράς του συστήματος.

Ο χρήστης έχει τη δυνατότητα να επιλέξει να απενεργοποιήσει τον μηχανισμό της δειγματοληψίας. Στη περίπτωση αυτή, η ακολουθία δεδομένων που παράγεται από τη μονάδα STU (Πίνακας 3) στο τέλος κάθε κύκλου αντιδράσεων αποθηκεύεται σε μια δομή δεδομένων (First-In First Out - FIFO). Στη FIFO αποθηκεύεται μόνο η πληροφορία που αφορά τον δείκτη μ της νικητήριας αντίδρασης R_μ , το αναγνωστικό του νικητήριου νήματος ($WTID$), το χρόνο t_μ , τις οχτώ ανανεωμένες τιμές των μοριακών πληθυσμών, τον χρόνο t_{sim} και τον μετρητή των αντιδράσεων. Έπειτα, οι συνολικά 14 λέξεις των 32-bit που είναι αποθηκευμένες στη FIFO μεταδίδονται στο δίκτυο υπό τον έλεγχο της FSM D.

Η μηχανισμός της δειγματοληψίας μπορεί να απενεργοποιηθεί και δυναμικά εάν κριθεί σκόπιμο από τη μονάδα OIS. Κάτι τέτοιο μπορεί να συμβεί στη περίπτωση όπου τα χρονικά βήματα τ_μ της προσομοίωσης που παράγονται από τη μονάδα FMU είναι μεγαλύτερα από την προκαθορισμένη από τον χρήστη περίοδο δειγματοληψίας T_{sam} ($\tau_\mu > T_{sam}$). Μια τέτοια περίπτωση συνεπάγεται ότι σε κάθε κύκλο αντίδρασης αυξάνεται και χρόνος προσομοίωσης t_{sim} αλλά και ο χρόνος δειγματοληψίας t_{sam} με αποτέλεσμα να μην υφίσταται πλέον

δειγματοληψία των αποτελεσμάτων. Δεδομένου των αρκετών κύκλων ρολογιού που απαιτούνται για την ολοκλήρωση των φάσεων του μηχανισμού δειγματοληψίας (αντιγραφή του πίνακα ST, διάβασμα του πίνακα ST, αποστολή δεδομένων) το σύστημα θα πρέπει να περιμένει προτού προχωρήσει στην έναρξη ενός νέου κύκλου αντιδράσεων. Αντιθέτως, για την αποστολή της ακολουθίας δεδομένων των 14 λέξεων που προαναφέρθηκε απαιτούνται λιγότεροι κύκλοι ρολογιού. Αναγνωρίζοντας λοιπόν μια τέτοια κατάσταση ($\tau_{\mu} > T_{sam}$) η μονάδα OIS είναι σε θέση να απενεργοποιήσει δυναμικά το μηχανισμό δειγματοληψίας γλυτώνοντας το σύστημα από το κόστος των επιπλέον κύκλων ρολογιού που απαιτούνται για την ολοκλήρωση του.

4.4.7 Αναπαράσταση Αποτελεσμάτων

Μέχρι τώρα έχουν αναφερθεί οι μηχανισμοί και οι τρόποι αποστολής δεδομένων. Στη συνέχεια παρουσιάζεται η δομή των πακέτων που στέλνονται στο δίκτυο από τους FRM πυρήνες. Στην Ενότητα 3 αναφέρθηκε ότι τα αποτελέσματα του SoC καταλήγουν στο Host PC για επεξεργασία. Τα δεδομένα αυτά αποθηκεύονται σειριακά σε ένα μοναδικό binary αρχείο στο οποίο πραγματοποιούνται εγγραφές από όλους τους πυρήνες του SoC. Προκειμένου να είναι εφικτός ο διαχωρισμός των τύπων των δεδομένων χρησιμοποιούνται συγκεκριμένες ετικέτες (labels) των 32-bit για κάθε τύπο. Στον πίνακα που ακολουθεί δίνονται οι ετικέτες που χρησιμοποιούνται κατά τη μεταφορά δεδομένων από τον FRM πυρήνα προς το Host PC.

Πίνακας 4: Τύποι και Ετικέτες Δεδομένων

Τύποι και Ετικέτες Δεδομένων		
Όνομα Ετικέτας	Τιμή Ετικέτας (Hex)	Περιγραφή
CORE ID LABEL	FFFFCCCC	Αναγνωριστικό (ID) FRM πυρήνα, έναρξη πακέτου δεδομένων
SP DATA LABEL	FFFFFFFF	Ετικέτα έναρξης δεδομένων μηχανισμού δειγματοληψίας
RC DATA LABEL	FFFFE000	Ετικέτα έναρξης δεδομένων ενός κύκλου αντιδράσεων
EOR LABEL	FFFFD000	Ετικέτα τερματισμού μιας στοχαστικής επανάληψης (repetition) προσομοίωσης

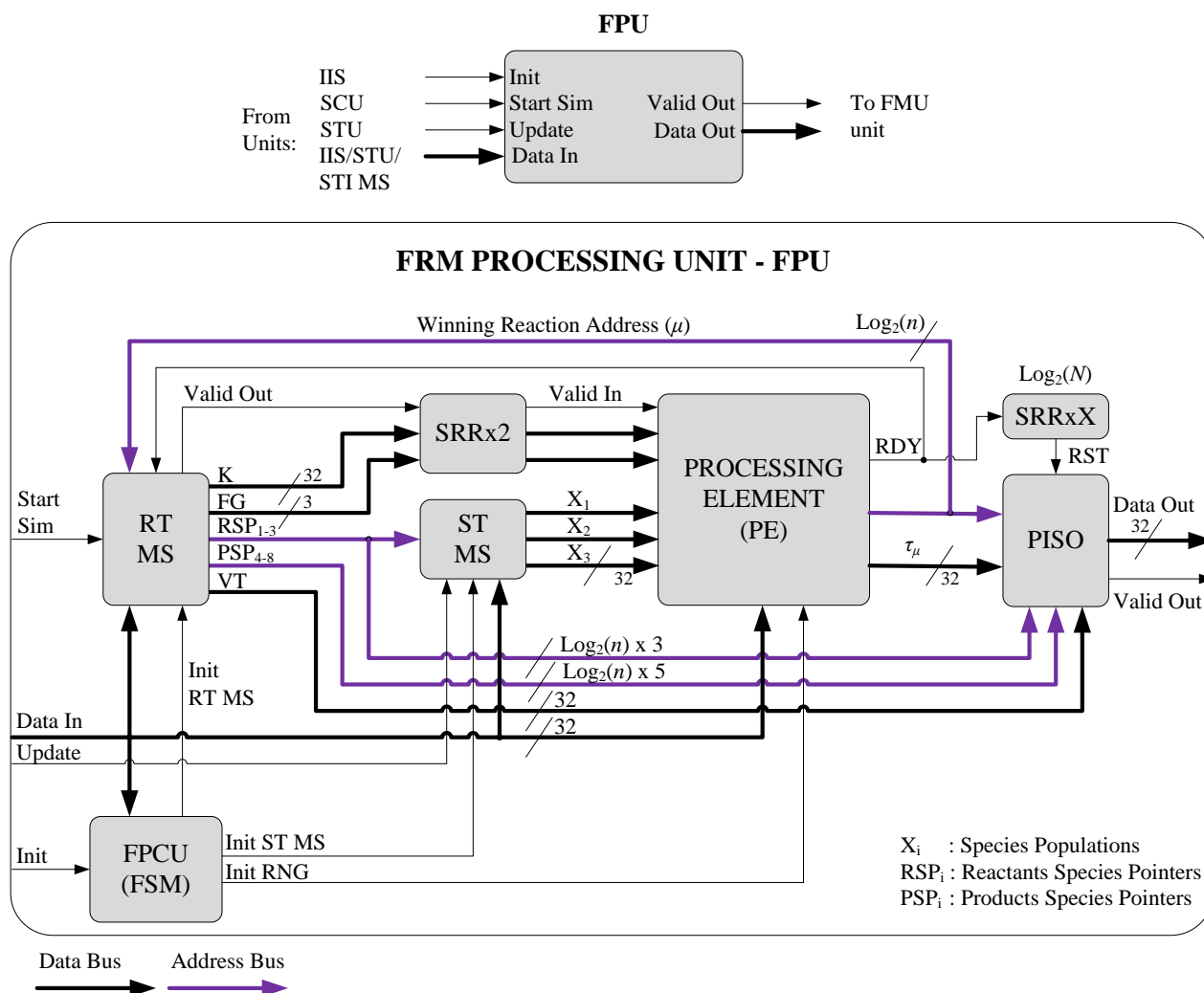
Η ετικέτα CORE ID LABEL υποδηλώνει την αρχή ενός πακέτου δεδομένων ενώ η αμέσως επόμενη λέξη που έπεται είναι το αναγνωριστικό (ID) του πυρήνα από τον οποίο προέκυψαν τα δεδομένα που ακολουθούν. Μετά από κάθε αναγνωριστικό στέλνεται η ετικέτα SP DATA LABEL ή ετικέτα RC DATA LABEL. Η ετικέτα SP DATA LABEL υποδηλώνει ότι τα δεδομένα που ακολουθούν έχουν προκύψει από το μηχανισμό δειγματοληψίας (Sampling Period - SP). Η αμέσως επόμενη λέξη που στέλνεται μετά από την ετικέτα SP DATA LABEL είναι η περίοδος δειγματοληψίας (t_{sam}) από την οποία προέκυψαν τα δεδομένα που ακολουθούν. Μετά από τη περίοδο δειγματοληψίας στέλνονται ζευγάρια δεικτών και τιμών των πληθυσμών που μεταβλήθηκαν κατά τη τελευταία περίοδο δειγματοληψίας. Η πρώτη λέξη του ζευγαριού αντιστοιχεί σε έναν δείκτη 32-bit ακεραίας αριθμητικής που ταυτοποιεί ένα μοριακό είδος ενώ η δεύτερη λέξη του ζευγαριού αντιστοιχεί στο πληθυσμό του μοριακού είδους (χρησιμοποιείται SPFP αριθμητική). Η ετικέτα RC DATA LABEL (Reaction Cycle - RC) υποδηλώνει ότι τα δεδομένα που ακολουθούν δεν έχουν προκύψει από το μηχανισμό δειγματοληψίας αλλά είναι μέρος της ακολουθία που παράγει η μονάδα STU. Η ακολουθία αυτή περιέχει τα εξής δεδομένα: τον δείκτη μ της νικητήριας αντίδρασης R_{μ} , τον χρόνο t_{μ} της νικητήριας αντίδρασης, τις οχτώ ανανεωμένες τιμές των μοριακών πληθυσμών, τους χρόνους t_{sim} και t_{sam} της προσομοίωσης και τον

μετρητή των αντιδράσεων. Μόλις ολοκληρωθεί μια στοχαστική επανάληψη προσομοίωσης στέλνεται η ετικέτα EOR LABEL (End of Repetition - EOR) ενώ οι αμέσως επόμενες δύο λέξεις που στέλνονται είναι το ψηλό και το χαμηλό κομμάτι του 64-bit μετρητή που καταγράφει τον συνολικό αριθμό αντιδράσεων που πραγματοποιήθηκαν κατά τη διάρκεια της προσομοίωσης. Η επιλογή των δεκαεξαδικών τιμών των ετικετών έγινε με τέτοιο τρόπο ούτως ώστε στα πλαίσια του προτύπου IEEE 754 [40] οι τιμές αυτές να ορίζονται ως αδιευκρίνιστες τιμές ή τιμές που δεν μπορούν να αναπαρασταθούν (Not a Number -NaN).



5. ΑΡΧΙΤΕΚΤΟΝΙΚΗ FRM ΕΠΕΞΕΡΓΑΣΤΙΚΗΣ ΜΟΝΑΔΑΣ

Στη προηγούμενη ενότητα παρουσιάστηκε η αρχιτεκτονική του FRM SoC και η αρχιτεκτονική του FRM πυρήνα. Σε αυτή την ενότητα θα παρουσιαστεί η αρχιτεκτονική της FRM επεξεργαστικής μονάδας (FRM Processing Unit - FPU). Στη συνέχεια ακολουθεί σχηματικό αρχιτεκτονικής της μονάδας FPU.



Εικόνα 15: Αρχιτεκτονική της μονάδας FPU

Η FPU μονάδα είναι υπεύθυνη για τον υπολογισμό του χρόνου τ_μ ενός υποσυνόλου αντιδράσεων (m/N) του βιομοντέλου και την εύρεση της αντίδρασης (R_μ) από την οποία προέκυψε ο χρόνος αυτός. Οι κύριες υπομονάδες της μονάδας FPU είναι τα συστήματα μνήμης RT MS και ST MS για την αποθήκευση των πινάκων RT και ST αντίστοιχα και μία μονάδα (Processing Element - PE) για τον υπολογισμό του χρόνου τ_μ . Σε κάθε κύκλο αντίδρασης (reaction cycle) αναλαμβάνει να υπολογίσει όλους τους χρόνους τ_j του υποσυνόλου (m/N) που της έχει ανατεθεί. Παράλληλα καθώς υπολογίζονται οι χρόνοι τ_j στην FPU μονάδα, συγκερατείται σε

έναν καταχωρητή της μονάδας ο μικρότερος χρόνος που έχει υπολογιστεί προκειμένου στο τέλος της διαδικασίας να προκύψει ο χρόνος τ_μ .

Επιπλέον η αρχιτεκτονική της μονάδας FPU αποτελείται μια μονάδα ελέγχου (FPCU) για την αρχικοποίηση των πινάκων RT και ST, μια μονάδα παράλληλης εισόδου – σειριακής εξόδου (Parallel-in – Serial-out - PISO) για την πολυπλεξία του διαύλου εξόδου (Data Out) καθώς και από κυκλώματα ολίσθησης (SRR).

Η FPU έχει δύο τύπους μνήμης. Η μνήμη RT βρίσκεται εντός της μονάδας RT MS και σε αυτή αποθηκεύεται ολόκληρος ή τμήματα (m/N) του πίνακα RT ανάλογα από τον αριθμό των FPU που διαθέτει ο FRM πυρήνας. Η μνήμη αυτή μπορεί να θεωρηθεί ως η Instruction Memory της FPU. Η μνήμη ST βρίσκεται εντός της μονάδας ST MS και σε αυτή αποθηκεύεται ο ολόκληρος πίνακας ST ανεξάρτητα από τον αριθμό των FPU. Η μνήμη αυτή μπορεί να θεωρηθεί ως η Data Memory της FPU.

Η μονάδα PISO χρειάζεται προκειμένου να πολυπλεχθούν όλα τα δεδομένα που παράγονται ή είναι αποθηκευμένα στη μονάδα FPU και είναι απαραίτητα στις μονάδες FMU και STU του FRM πυρήνα. Τα δεδομένα αυτά είναι τα εξής: ο χρόνος τ_μ ($\text{Min } \tau_j$) που υπολογίζεται από τη μονάδα PE, η κωδικοποιημένη στοιχειομετρία (VT) της νικητήριας αντίδρασης R_μ , οι οχτώ δείκτες των πληθυσμών των μοριακών ειδών (Species Pointers - SP_i) που συμμετέχουν στην αντίδραση R_μ καθώς και ο δείκτης μ της αντίδρασης R_μ (Winning Reaction Address). Συνεπώς, η PISO μονάδα είναι υπεύθυνη για τη σειριακή μετάδοση της προαναφερθείσας ακολουθίας δεδομένων μέσω ενός και μόνο 32-bit διαύλου.

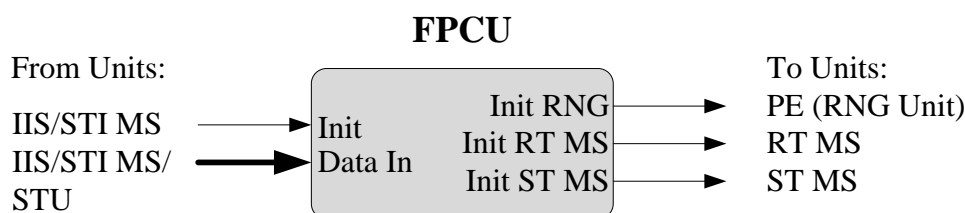
Στην Εικόνα 15 φαίνονται τρία κυκλώματα ολίσθησης μήκους δύο θέσεων (SRRx2) και ένα κύκλωμα ολίσθησης μήκους $\text{Log}_2(N)$ θέσεων. Τα κυκλώματα ολίσθησης δύο θέσεων απαιτούνται για το συγχρονισμό των σημάτων *RT Valid Out/PE Valid In*, *K* και *FG* με τους πληθυσμούς των μοριακών ειδών που προσκομίζονται από το σύστημα μνήμης ST MS το οποίο έχει καθυστέρηση λόγω διοχέτευσης δύο κύκλους ρολογιού. Το κύκλωμα ολίσθησης μήκους $\text{Log}_2(N)$ θέσεων απαιτείται για την αρχικοποίηση της μονάδας PISO Unit αφού πρώτα εντοπιστεί από τη μονάδα CMP TREE της FMU ο συνολικά μικρότερος χρόνος τ_μ της νικητήριας αντίδρασης R_μ . Μέχρι τότε η μονάδα PISO τροφοδοτεί στην έξοδο της μόνο τη τιμή τ_μ που υπολογίστηκε από τη μονάδα FPU και όχι την ακολουθία δεδομένων που προαναφέρθηκε.

Η λειτουργία της FPU μονάδας συμβαδίζει με τις τρεις φάσεις *Initialization - 1*, *Computation (τ_μ) - 2* και *Update Sim State - 3* του FRM πυρήνα που περιγράφηκαν στην Ενότητα 4.4. Στην FPU μονάδα και τα δύο στάδια της φάσης (1) ελέγχονται από την μονάδα ελέγχου FPCU μέσω των σημάτων *Init RT MS*, *Init ST MS* και *Init RNG*. Η έναρξη της φάσης (2) σηματοδοτείται μέσω του σήματος *Start Sim* το οποίο οδηγείται στην μονάδα RT MS προκειμένου να ξεκινήσει η ανάγνωση του πίνακα RT. Μόλις η μονάδα RT MS τροφοδοτήσει στην έξοδο της έγκυρα δεδομένα θα θέσει στη λογική κατάσταση (1) το σήμα *RT MS Valid Out* το οποίο μετά από δύο κύκλους ρολογιού, λόγω του ολισθητή, θα ενημερώσει τη PE μονάδα ότι θα πρέπει να ξεκινήσει τη λειτουργία της. Η φάση (3) σηματοδοτείται μέσω του σήματος *Update* το οποίο οδηγείται στην μονάδα ST MS προκειμένου να ξεκινήσει η ανανέωση των πληθυσμών των μοριακών ειδών.

Τέλος θα πρέπει να αναφερθεί ότι η φάση (2) στην FPU μονάδα αποτελείται από δύο στάδια. Το πρώτο στάδιο περιλαμβάνει τον υπολογισμό του χρόνου τ_μ και την εύρεση του δείκτη μ της νικητήριας αντίδρασης R_μ . Το δεύτερο στάδιο περιλαμβάνει την εύρεση της νικητήριας αντίδρασης R_μ στον πίνακα RT της μονάδας RT MS μέσω του δείκτη μ και την διοχέτευση των πεδίων της (*βλ.* Ενότητα 4.3) στην FMU μονάδα μέσω της PISO μονάδας.

5.1 FRM Processing Control Unit - FPCU

Για τον έλεγχο και την ολοκλήρωση της φάσης της αρχικοποίησης αναπτύχθηκε μονάδα ελέγχου (control unit) η οποία υλοποιήθηκε με τη χρήση μιας μηχανής FSM. Η FSM, κατά το στάδιο της αρχικοποίησης, μπορεί να παρομοιαστεί με τη λειτουργία ενός αποκωδικοποιητή ο οποίος δέχεται στην είσοδο του ένα σήμα (*Init*) και οδηγεί στην έξοδο τρία σήματα (*Init RNG*, *Init RT* και *Init ST*). Μέσω αυτών των σημάτων ελέγχου και με τη χρήση εσωτερικών μετρητών η FSM είναι σε θέση να αρχικοποιήσει τους πίνακες RT και ST και τη γεννήτρια ψευδοτυχαίων αριθμών. Για την ορθή οδήγηση των προαναφερθέντων σημάτων η FSM χρειάζεται να γνωρίζει πληροφορίες που σχετίζονται με το βιομοντέλο αλλά και με τις διαθέσιμες FPU του FRM πυρήνα. Τις πληροφορίες που σχετίζονται με το βιομοντέλο η FSM τις λαμβάνει μέσω του αρχείου παραμέτρων (PAR) κατά το στάδιο της αρχικοποίησης του FRM πυρήνα, ενώ τον αριθμό των διαθέσιμων FPU τον γνωρίζει μέσω Generic παραμέτρων. Στη συνέχεια ακολουθεί σχηματικό μπλοκ διαγράμματος με τα σήματα εισόδου και εξόδου της μονάδας FPCU.



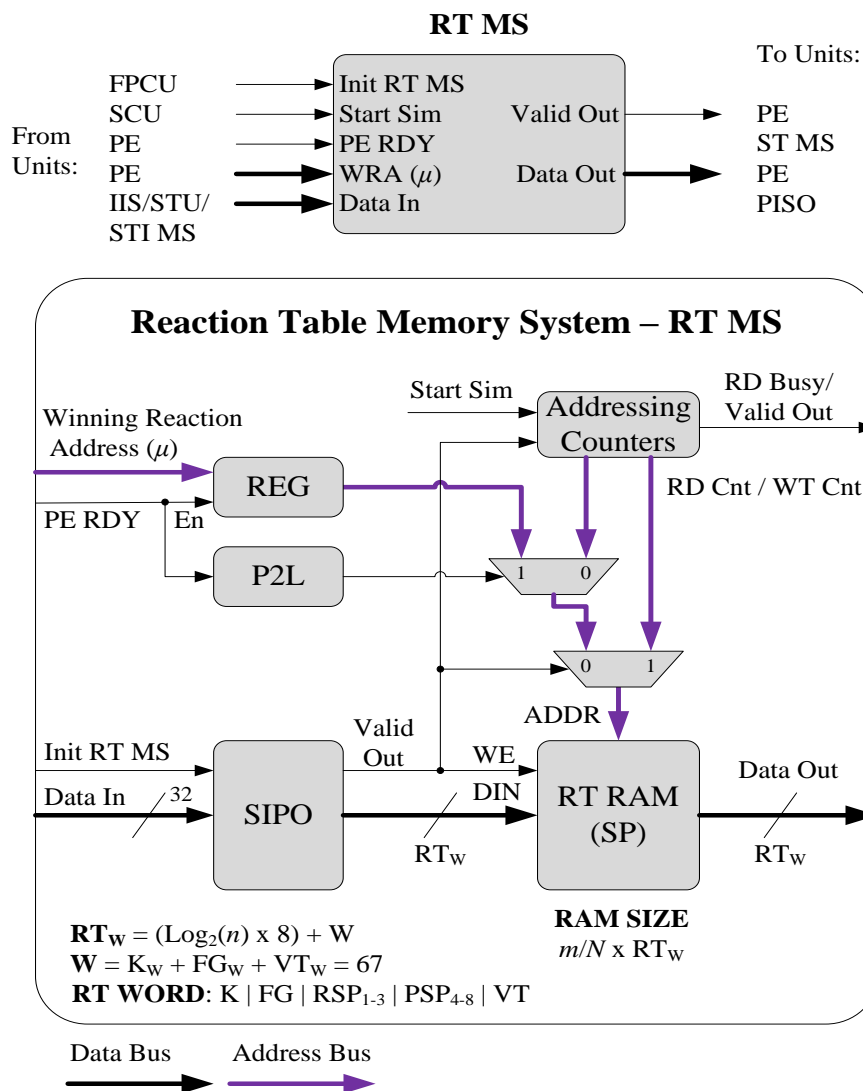
Εικόνα 16: Μπλοκ διάγραμμα της μονάδας FPCU

5.2 Reaction Table Memory System – RT MS

Η μονάδα Reaction Table Memory System – RT MS είναι υπεύθυνη για τις εξής λειτουργίες:

- Αποθήκευση του υποσυνόλου αντιδράσεων που αντιστοιχεί στην εκάστοτε FPU κατά το στάδιο της αρχικοποίησης του FRM πυρήνα
- Για το διάβασμα του υποσυνόλου αντιδράσεων κατά το στάδιο υπολογισμού του χρόνου τ_μ
- Για την αναζήτηση της αντίδρασης R_μ (Winning Reaction) κατά το δεύτερο στάδιο της φάσης *Computation* (τ_μ) της μονάδας FPU μέσω του δείκτη μ (Winning Reaction Address)

Στην Εικόνα 17 δίνεται σχηματικό αρχιτεκτονικής της μονάδας RT MS. Για την αποθήκευση των δεδομένων χρησιμοποιείται μια μνήμη Single Port (SP) RAM (RT RAM) βάθους m/N λέξεων με παραμετρικό εύρος λέξης που καθορίζεται από τη σταθερά RT_w . Η σταθερά RT_w προκύπτει από το εύρος των πεδίων που απαρτίζουν κάθε λέξη του πίνακα RT. Σε κάθε λέξη της μνήμης RT RAM αποθηκεύονται τα εξής δεδομένα: ο ρυθμός αντίδρασης (K) 32-bit, η τάξης της αντίδρασης (FG) 3-bit, τρεις δείκτες (Reactant Species Pointers - RSP) προς τους πληθυσμούς των αντιδρώντων $\log_2(n) \times 3$ -bit, πέντε δείκτες (Product Species Pointers - PPP) προς τους πληθυσμούς των προϊόντων $\log_2(n) \times 5$ -bit και η στοιχειομετρία (VT) της αντίδρασης 32-bit. Από τα προαναφερθέντα δεδομένα, τα δεδομένα K, FG και VT έχουν σταθερό εύρος δίνοντας άθροισμα 67-bit. Αντιθέτως οι δείκτες RP και PP εξαρτώνται από τον αριθμό των μοριακών ειδών (n) του βιομοντέλου (generic parameters). Συνεπώς, η σταθερά RT_w προκύπτει από τη σχέση $(\log_2(n) \times 8) + 67$. Η σύνθεση της μνήμης RT RAM γίνεται με τη χρήση των διαθέσιμων BRAMs του εκάστοτε FPGA.



Εικόνα 17: Αρχιτεκτονική της μονάδας RT MS

Για την αποθήκευση των δεδομένων χρησιμοποιείται μια μνήμη Single Port (SP) RAM (RT RAM) βάθους m/N λέξεων με παραμετρικό εύρος λέξης που καθορίζεται από τη σταθερά RT_w . Η σταθερά RT_w προκύπτει από το εύρος των πεδίων που απαρτίζουν κάθε λέξη του πίνακα RT. Σε κάθε λέξη της μνήμης RT RAM αποθηκεύονται τα εξής δεδομένα: ο ρυθμός αντίδρασης (K) 32-bit, η τάξης της αντίδρασης (FG) 3-bit, τρεις δείκτες (Reactant Species Pointers - RSP) προς τους πληθυσμούς των αντιδρώντων $\text{Log}_2(n) \times 3$ -bit, πέντε δείκτες (Product Species Pointers - PPP) προς τους πληθυσμούς των προϊόντων $\text{Log}_2(n) \times 5$ -bit και η στοιχειομετρία (VT) της αντίδρασης 32-bit. Από τα προαναφερθέντα δεδομένα, τα δεδομένα K, FG και VT έχουν σταθερό εύρος δίνοντας άθροισμα 67-bit. Αντιθέτως οι δείκτες RP και PP εξαρτούνται από τον αριθμό των μοριακών ειδών (n) του βιομοντέλου (generic parameters). Συνεπώς, η σταθερά RT_w προκύπτει από τη σχέση $(\text{Log}_2(n) \times 8) + 67$. Η σύνθεση της μνήμης RT RAM γίνεται με τη χρήση των διαθέσιμων BRAMs του εκάστοτε FPGA.

Στο Host PC οι αντιδράσεις της μνήμης RT RAM είναι αποθηκευμένες σε ένα δυαδικό αρχείο όπου κάθε αντίδραση αποτελείται από επτά λέξεις των 32-bit λόγω της 32-bit αρχιτεκτονικής του συστήματος. Κατά το στάδιο της αρχικοποίησης, από το Host PC στέλνονται οι λέξεις αυτές και από τη πλευρά του FRM SoC συλλέγονται και σταδιακά δημιουργούνται οι αντιδράσεις/λέξεις, εύρους RT_w -bit, της μνήμης RT RAM. Υπεύθυνη για αυτή τη λειτουργία είναι μια μονάδα Σειριακής-εισόδου - Παράλληλης-εξόδου (Serial-in Parallel-out – SIPO). Μόλις η SIPO μονάδα σχηματίσει μια αντίδραση θέτει το σήμα της *Valid Out* προκειμένου να πραγματοποιηθεί μια εγγραφή στη μνήμη RT RAM και να αυξηθεί ο μετρητής εγγραφών (WT Cnt) που χρησιμοποιείται για τη διευθυνσιοδότηση της μνήμης κατά το στάδιο της αρχικοποίησης.

Η έναρξη της φάσης υπολογισμού του χρόνου τ_μ σηματοδοτείται από το σήμα *Start Sim* που οδηγείται από τη μονάδα ελέγχου SCU. Το σήμα *Start Sim* ενεργοποιεί τον μετρητή RD Cnt που χρησιμοποιείται για τη διευθυνσιοδότηση της μνήμης RT RAM κατά το στάδιο αυτό. Ως εκ τούτου, κατά το στάδιο αυτό διαβάζονται διαδοχικά όλες οι αντιδράσεις που είναι αποθηκευμένες στη μνήμη RT RAM οι οποίες στη συνέχεια τροφοδοτούνται στην έξοδο της μονάδας RT MS. Οι υπόλοιπες υπομονάδες της FPU ενημερώνονται ότι μεταδίδονται έγκυρα δεδομένα από τη μονάδα RT MS μέσω του σήματος *RD Busy/Valid Out* το οποίο μεταβαίνει για το λόγο αυτό στη λογική κατάσταση (1). Θα πρέπει να αναφερθεί ότι η πρώτη αντίδραση θα τροφοδοτηθεί μετά από δύο κύκλους ρολογιού στην έξοδο της μονάδας RT MS. Ο λόγος της καθυστέρησης αυτής είναι ο ένας κύκλος ρολογιού που απαιτείται για το διάβασμα της μνήμης συν ένα κύκλο για το καταχωρητή εξόδου της μνήμης. Ο καταχωρητής εξόδου της μνήμης επιλέχθηκε να χρησιμοποιηθεί καθώς παρατηρήθηκαν κατά το στάδιο της σχεδίασης αυξημένα συνδυαστικά μονοπάτια από το μνήμη προς άλλα κυκλώματα.

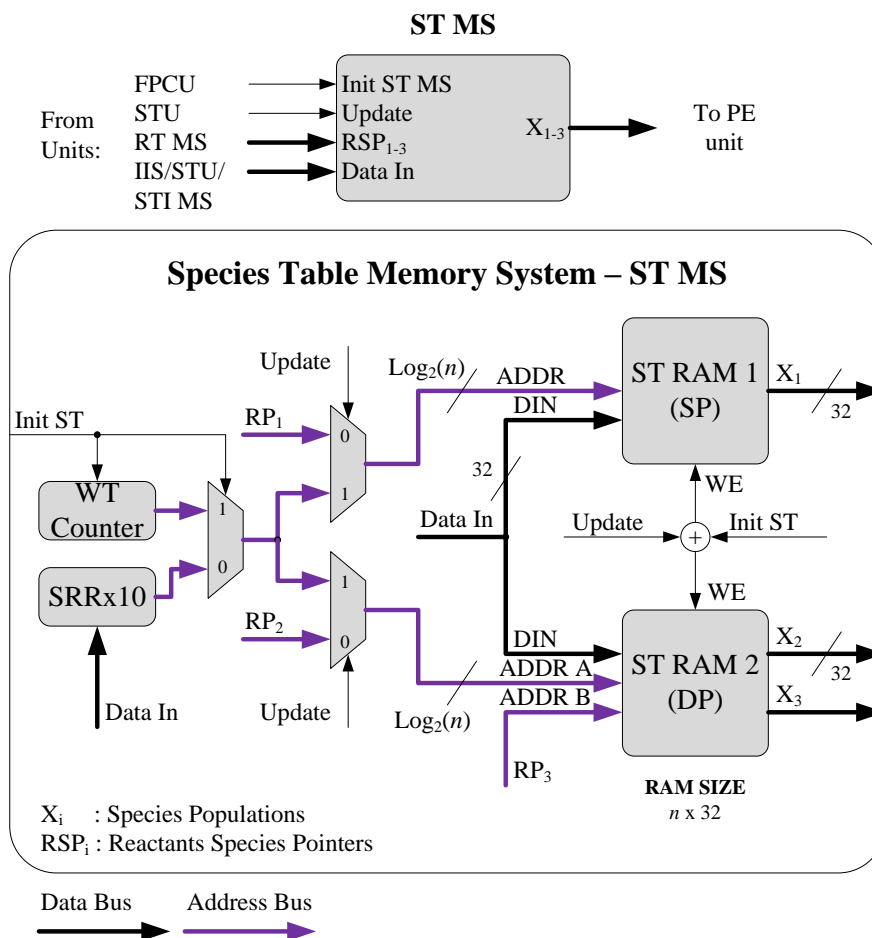
Με την ολοκλήρωση του υπολογισμού του χρόνου τ_μ θα πρέπει να βρεθεί η αντίδραση R_μ (Winning Reaction). Αυτό γίνεται με τη χρήση του δείκτη μ της αντίδρασης R_μ . Η μονάδα PE η οποία υπολογίζει τον χρόνο τ_μ ενημερώνει τη μονάδα RT MS ότι υπολογίστηκε ο χρόνος τ_μ μέσω ενός παλμού. Ο παλμός αυτός τροφοδοτείται στην είσοδο *PE RDY* της μονάδας RT MS. Το σήμα *PE RDY* ενεργοποιεί έναν καταχωρητή προκειμένου να αποθηκευθεί ο δείκτης της R_μ . Στη συνέχεια ο καταχωρητής αυτός θα χρησιμοποιηθεί για την διευθυνσιοδότηση της μνήμης RT RAM, καθ' όλη τη διάρκεια του δεύτερου σταδίου της φάσης *Computation* (τ_μ). Ο παλμός του σήματος *PE RDY* πρέπει να μετατραπεί σε ένα νέο σήμα με "σταθερή λογική κατάσταση". Το νέο αυτό σήμα θα πρέπει να διατηρεί τη λογική κατάσταση (1) για όσο χρονικό διάστημα διαρκεί το στάδιο αυτό προκειμένου να οδηγήσει κατάλληλα τον άνω πολυπλέκτη της Εικόνα 17. Τη μετατροπή αυτή αναλαμβάνει να πραγματοποιήσει μια μονάδα μετατροπής παλμών σε σταθερή λογική κατάσταση για προκαθορισμένο χρονικό διάστημα (Pulse-to-Level – P2L).

5.3 Species Table Memory System – ST MS

Η μονάδα Species Table Memory System – ST MS είναι υπεύθυνη για τις εξής λειτουργίες:

- Αποθήκευση των πληθυσμών των μοριακών ειδών του πίνακα ST κατά τη διάρκεια της φάσης *Initialization* του FRM πυρήνα
- Προσκόμιση των πληθυσμών των μοριακών ειδών στη μονάδα PE κατά τη διάρκεια της φάσης *Computation* (τ_μ)
- Ανανέωση των πληθυσμών των μοριακών ειδών κατά τη διάρκεια της φάσης *Update Sim State*

Στη συνέχεια ακολουθεί σχηματικό αρχιτεκτονικής της μονάδας ST MS.



Εικόνα 18: Αρχιτεκτονική της μονάδας ST MS

Για την αποθήκευση των δεδομένων χρησιμοποιούνται δύο μνήμες RAM βάθους n λέξεων όπου η κάθε λέξη έχει εύρος 32-bit λόγω της απλής ακρίβειας που χρησιμοποιείται για την αναπαράσταση των πληθυσμών των μοριακών ειδών. Η μνήμη ST RAM 1 είναι τύπου Single Port (SP) ενώ η μνήμη ST RAM 2 είναι τύπου Dual Port (DP). Και στις δύο μνήμες αποθηκεύεται ολόκληρος ο πίνακας ST, άσχετα με τον αριθμό των διαθέσιμων FPU's (N), ενώ και στις δύο μνήμες γράφονται ακριβώς τα ίδια δεδομένα. Για το λόγο αυτό, έχουν κοινά σήματα για την είσοδο δεδομένων (DIN) και κοινά σήματα ενεργοποίησης εγγραφών (WE). Ο λόγος που η μονάδα ST MS χρησιμοποιεί δύο μνήμες RAM σχετίζεται με τη τάξη των αντιδράσεων που υποστηρίζει το σύστημα FRM SoC. Στην Ενότητα 3, αναφέρθηκε ότι το σύστημα έχει σχεδιαστεί για να υποστηρίζει αντιδράσεις που περιέχουν μέχρι 3 διαφορετικά μοριακά είδη. Προκειμένου η μονάδα υπολογισμού του χρόνου τ_μ να μπορεί να εξετάζει μια αντίδραση ανά κύκλο ρολογιού απαιτείται να τροφοδοτούνται σε αυτή ταυτόχρονα τρεις πληθυσμοί μοριακών ειδών. Χρησιμοποιώντας λοιπόν δύο μνήμες RAM, μια SP και μια DP, η μονάδα ST MS μπορεί να προσκομίσει ταυτόχρονα τρεις πληθυσμούς μοριακών ειδών (X_{1-3}) στη μονάδα PE. Η σύνθεση των μνημών ST RAM γίνεται με τη χρήση των διαθέσιμων BRAMs του εκάστοτε FPGA.

Κατά τη διάρκεια της φάσης *Initialization* στέλνονται σταδιακά από τη μονάδα STI MS οι 32-bit τιμές του πίνακα ST οι οποίες και αποθηκεύονται στις μνήμες ST RAM. Η αρχικοποίηση σηματοδοτείται μέσω του σήματος *Init ST*. Το σήμα *Init ST* ενεργοποιεί τις εγγραφές στις μνήμες ST RAM ενώ αυξάνει και κατά ένα το μετρητή εγγραφών (WT Cnt) που χρησιμοποιείται για τη διευθυνσιοδότηση των μνημών στη φάση της αρχικοποίησης.

Κατά τη διάρκεια της φάσης Computation (τ_μ) λαμβάνονται από τη μονάδα RT MS οι δείκτες RSP₁₋₃ των μοριακών αντιδρώντων οι οποίοι χρησιμοποιούνται για τη διευθυνσιοδότηση των μνημών στη φάση αυτή. Όπως φαίνεται και από την αρχιτεκτονική της μονάδας ST MS στην Εικόνα 18 κάτι τέτοιο είναι εφικτό λόγω της κατάστασης των πολυπλεκτών σε αυτή τη φάση εφόσον τα σήματα *Init ST* και *Update ST* βρίσκονται στη λογική τιμή (0).

Στη μονάδα ST MS η φάση *Update Sim State* σηματοδοτείται μέσω του σήματος *Update* το οποίο οδηγείται από τη μονάδα STU κάθε φορά που μεταδίδει έγκυρα δεδομένα στο κύριο δίαυλο του FRM πυρήνα. Σύμφωνα με τον Πίνακα 3 η μονάδα STU αρχικά μεταδίδει τους δείκτες των μοριακών ειδών (SP_i) της νικητήριας αντίδρασης R_μ σε κάθε κύκλο προσομοίωσης του αλγορίθμου. Έπειτα μεταδίδει το δείκτη μ της αντίδρασης R_μ , το αναγνωριστικό του νικητήριου νήματος (WT_{ID}), και μετά συνεχίζει με τη μετάδοση των ανανεωμένων πληθυσμών S_i και του μετρητή των αντιδράσεων.

Προκειμένου να δημιουργηθούν ζευγάρια δεικτών και τιμών (SP_i, X_i) ούτως ώστε να γίνουν σωστά οι εγγραφές στις μνήμες ST MS χρησιμοποιείται μια μονάδα ολισθησης δέκα θέσεων (SRRx10). Οι δέκα θέσεις του ολισθητή προκύπτουν από τις δέκα λέξεις που προαναφέρθηκαν και μεσολαβούν μέχρι να ξεκινήσουν να μεταδίδονται οι τιμές X_i από τη μονάδα STU. Συνεπώς, η έξοδος του ολισθητή μετά από δέκα κύκλους ρολογιού θα αρχίσει να τροφοδοτεί τη μονάδα ST MS με τους δείκτες SP_i τη στιγμή ακριβώς που θα έχει φτάσει η πρώτη τιμή X_i. Εφόσον στη φάση αυτή, το σήμα *Init ST* βρίσκεται στη λογική κατάσταση (0) και το σήμα *Update ST* βρίσκεται στη λογική κατάσταση (1), όπως φαίνεται και στην Εικόνα 18, η διευθυνσιοδότηση των μνημών γίνεται μέσω των δεικτών SP_i. Ως εκ τούτου σχηματίζονται ζευγάρια SP_i, X_i και σε οχτώ κύκλους ρολογιού θα ολοκληρωθεί η ανανέωση των πινάκων ST με τους νέους πληθυσμούς.

5.4 Processing Element - PE

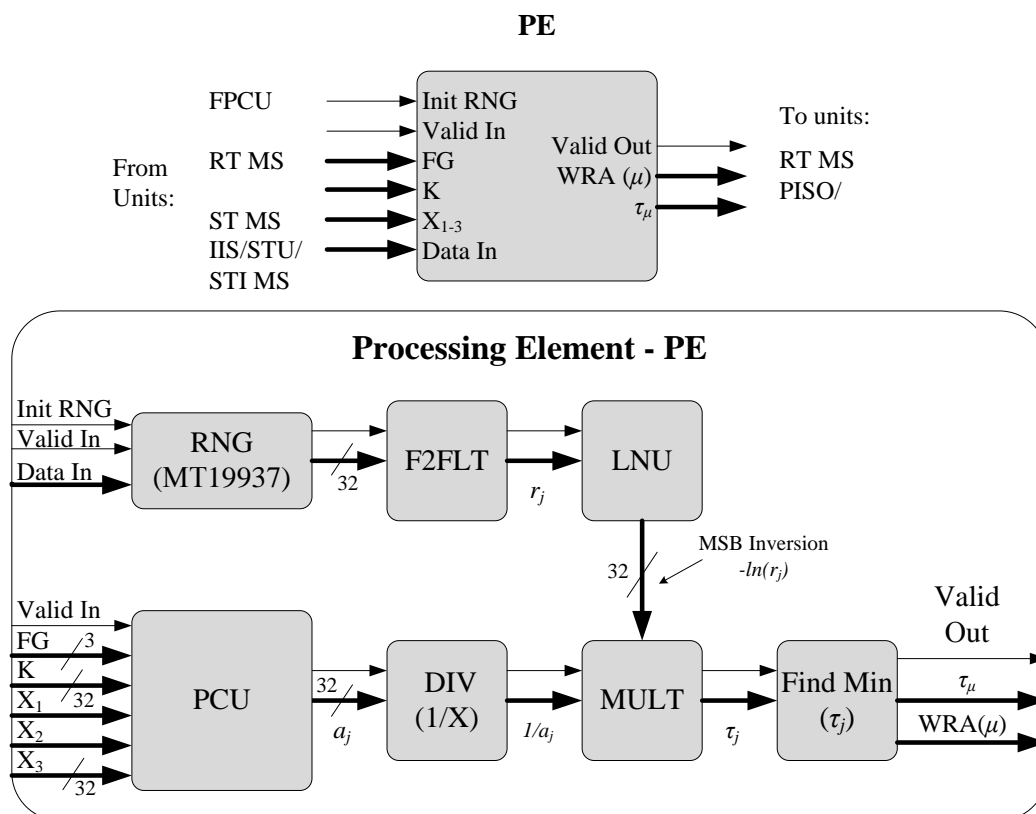
Μέχρι τώρα έχουν παρουσιαστεί όλες οι βοηθητικές μονάδες της FPU. Στη συνέχεια ακολουθεί η αναλυτική περιγραφή της μονάδας που είναι υπεύθυνη για τον υπολογισμό του χρόνου τ_μ . Στη συνέχεια δίνεται η σχέση από την οποία υπολογίζονται τα χρονικά βήματα τ_j σύμφωνα με τον αλγόριθμο FRM-SSA που παρουσιάστηκε στην Ενότητα 2.

(2)

$$\tau_j = \frac{1}{\alpha_j(x)} [-\ln(r_j)]$$

Ο τ_μ είναι ο μικρότερος τ_j χρόνος που υπολογίστηκε για όλο το σύνολο των αντιδράσεων του πίνακα RT που έχει ανατεθεί στην εκάστοτε FPU. Εκτός από τον χρόνο τ_μ στην έξοδο της μονάδας PE τροφοδοτείται και ο

δείκτης μ της αντίδρασης R_μ από την οποία προέκυψε ο χρόνος τ_μ . Στη συνέχεια ακολουθεί σχηματικό αρχιτεκτονικής της μονάδας PE.



Εικόνα 19: Αρχιτεκτονική της μονάδας PE

Η μονάδα PE αποτελείται κυρίως από Floating Point αριθμητικές μονάδες των 32-bit με πολλαπλά επίπεδα διασωλήνωσης πλήρους διοχέτευσης για υψηλότερες επιδόσεις. Για τις αριθμητικές πράξεις πολλαπλασιασμού, διαίρεσης, πρόσθεσης και υπολογισμού του φυσικού λογάριθμου έχουν χρησιμοποιηθεί μονάδες από τη βιβλιοθήκη Core Generation της Xilinx [54]. Οι διεπαφές αυτών των μονάδων στηρίζονται στο πρωτόκολλο AXIS4 [50]. Συνεπώς η δίοδος δεδομένων (datapath) της αρχιτεκτονικής του PE στηρίζεται σε σήματα τύπου Master/Slave του AXIS4. Αυτό σημαίνει ότι μόλις το σήμα *Valid in* της μονάδας PE λάβει τη λογική τιμή (1) θα ενεργοποιήσει τις μονάδες που το δέχονται ως είσοδο. Έπειτα, μόλις κάποια μονάδα παράγει έγκυρα δεδομένα θα ενεργοποιήσει με τη σειρά της την επόμενη μονάδα της δίοδου δεδομένων. Η συμπεριφορά αυτή καθιστά σχετικά εύκολη τη διαδικασία συγχρονισμού ανεξάρτητων μονάδων μιας δίοδου δεδομένων. Αποτέλεσμα αυτής της συμπεριφοράς είναι η κάθε η μονάδα να ενεργοποιείται μόνο για όσο χρειάζεται ενώ με τη σειρά της ενεργοποιεί τις υπόλοιπες μονάδες στις οποίες είναι συνδεδεμένη μόνο για το χρονικό διάστημα που παράγει έγκυρα δεδομένα.

Το πάνω μονοπάτι δεδομένων της Εικόνα 19 είναι υπεύθυνο για τον υπολογισμό του δεύτερου όρου της σχέσης (2). Αρχικά ενεργοποιείται, μέσω του σήματος *Valid In*, η γεννήτρια ψευδοτυχαίων αριθμών (Random Number Generation - RNG). Η RNG μονάδα παράγει ψευδοτυχαίους αριθμούς ανά ένα κύκλο ρολογιού για όσο χρονικό διάστημα το σήμα *Valid In* έχει τη λογική τιμή (1). Σύμφωνα με τον αλγόριθμο FRM-SSA, σε κάθε κύκλο

αντιδράσεων θα πρέπει να παραχθούν m ψευδοτυχαίοι αριθμοί. Εφόσον, κάθε PE μονάδα επεξεργάζεται τμήματα του πίνακα RT, ανάλογα με τη παράμετρο N , το σήμα *Valid In* θα παραμείνει στη λογική κατάσταση (1) για m/N κύκλους ρολογιού παράγοντας m/N ψευδοτυχαίους αριθμούς. Για τη παραγωγή των ψευδοτυχαίων αριθμών επιλέξαμε τον αλγόριθμο Mersenne Twister - MT19937 των Makoto Matsumoto και Takuji Nishimura [55]. Η απόφαση αυτή πάρθηκε για δύο βασικούς λόγους: Ο πρώτος λόγος είναι η ομοιόμορφη κατανομή που παρουσιάζουν οι παραγόμενοι αριθμοί του MT19937 αλγορίθμου κάτι το οποίο είναι απαίτηση του αλγορίθμου FRM-SSA για τους ψευδοτυχαίους μοναδιαίους αριθμούς r_j . Ο δεύτερος λόγος είναι η εξαιρετικά μεγάλη περίοδος που παρουσιάζει ο αλγόριθμος MT η οποία είναι της τάξεως $2^{19937}-1$. Υπό φυσιολογικές συνθήκες εκτέλεσης ενός πειράματος δεν αναμένεται να χρησιμοποιηθεί ποτέ ο ίδιος r_j ψευδοτυχαίος αριθμός. Εδώ θα πρέπει να αναφερθεί ότι ο αλγόριθμος MT19937 είναι η πρώτη επιλογή για την υλοποίηση γεννητριών ψευδοτυχαίων αριθμών για στοχαστικές μεθόδους προσομοίωσης. Ένα τέτοιο παράδειγμα είναι ο δημοφιλής προσομοιωτής Copasi που χρησιμοποιεί και αυτός τον αλγόριθμο MT19937. Για την FPGA υλοποίηση του αλγορίθμου MT19937 χρησιμοποιήσαμε τον IP πυρήνα που δίνεται στον ιστότοπο [56].

Για τη χρήση της μονάδας RNG-MT19937 χρειάστηκε να γίνουν δύο αναγκαίες τροποποιήσεις. Η πρώτη τροποποίηση αφορά τη συνδεσμολογία της μονάδας προκειμένου να παράγονται αριθμοί στο κλειστό διάστημα $[0, 1]$. Η δεύτερη τροποποίηση σχετίζεται με το γεγονός ότι οι αριθμοί που παράγονται είναι σταθερής υποδιαστολής (Fixed Point). Ωστόσο, οι υπόλοιπες μονάδες της δίοδου δεδομένων παράγουν αριθμούς κινητής υποδιαστολής (Floating Point). Για το λόγο αυτό, η έξοδος της μονάδας RNG-MT19937 τροφοδοτείται σε μια μονάδα μετατροπής αριθμών σταθερής υποδιαστολής σε αριθμούς κινητής υποδιαστολής απλή ακρίβειας. Στην Εικόνα 19 η μονάδα αυτή αναφέρεται ως F2FLT (Fixed to Float). Συνδέοντας λοιπόν στη γείωση το πιο σημαντικό ψηφίο (MSB) της εξόδου της μονάδας RNG-MT19937 παίρνουμε αριθμούς σταθερής υποδιαστολής στο κλειστό διάστημα $[0, 1]$ και τροφοδοτώντας τους αριθμούς αυτούς στη μονάδα F2FLT λαμβάνουμε τους ψευδοτυχαίους αριθμούς r_j κινητής υποδιαστολής. Στο σημείο αυτό θα πρέπει να αναφερθεί ότι η ακολουθία ψευδοτυχαίων αριθμών που παράγονται από την γεννήτρια MT19937 εξαρτάται από το διάνυσμα αρχικοποίησης της. Τροφοδοτώντας διαφορετικό διάνυσμα προκύπτει διαφορετική ακολουθία δίνοντας την δυνατότητα επαναχρησιμοποίησης της γεννήτριας στον FRM πυρήνα. Συνεπώς κάθε FPU της μονάδας FMU χρησιμοποιεί διαφορετική ακολουθία ψευδοτυχαίων αριθμών για τον υπολογισμό του χρόνου τ_i εφόσον η αντίστοιχη γεννήτρια αρχικοποιείται με διαφορετικό διάνυσμα. Η αρχικοποίηση της γεννήτριας MT19937 γίνεται στη φάση *Initialization* της μονάδας FPU και ελέγχεται από την μονάδα FPCU μέσω του σήματος *Init RNG*.

Οι ψευδοτυχαίοι αριθμοί που παράγονται από την μονάδα MT19937, στο κλειστό διάστημα $[0, 1]$, τροφοδοτούνται σε μια μονάδα υπολογισμού του φυσικού λογάριθμου (Ln Unit - LNU) αφού πρώτα έχουν μετατραπεί σε Floating Point αριθμητική μέσω της μονάδας F2FLT. Έπειτα, αντιστρέφοντας το πιο σημαντικό ψηφίο (MSB) της εξόδου της μονάδας LNU λαμβάνουμε τον δεύτερο όρο της σχέσης (2) $[-\ln(r_j)]$. Η αντιστροφή αυτή προκύπτει από το μείον της σχέσης (2). Κάτι το οποίο είναι απαραίτητο καθώς ο φυσικός λογάριθμος αριθμών μεταξύ μηδέν και ένα είναι αριθμός μικρότερος του μηδενός.

Το κάτω σωληνωμένο μονοπάτι της μονάδας PE είναι υπεύθυνο για τον υπολογισμό του πρώτου όρου της σχέσης (2). Η μονάδα Propensity Calculation Unit – PCU είναι υπεύθυνη για τον υπολογισμό των τάσεων (propensities) a_j των αντιδράσεων ενός βιομοντέλου. Μετά από μερικούς κύκλους ρολογιού, λόγω καθυστερήσεων της διοχέτευσης (latency), η μονάδα PCU θα αρχίσει να τροφοδοτεί σε κάθε κύκλο ρολογιού τη μονάδα DIV με τάσεις a_j . Στη συνέχεια, οι τάσεις a_j τροφοδοτούνται σε μια μονάδα διαίρεσης (DIV 1/X) για

τον υπολογισμό του λόγου ($1/a_j$) της σχέσης (2). Έπειτα, τα δεδομένα του επάνω μονοπατιού και του κάτω μονοπατιού, συγχρονισμένα τροφοδοτούνται σε μια μονάδα πολλαπλασιασμού (MULT) από την οποία θα προκύψουν οι χρόνοι t_j όλων των αντιδράσεων του τμήματος του πίνακα RT που διαχειρίζεται η εκάστοτε FPU.

Όλοι οι δυνητικοί χρόνοι t_j των αντιδράσεων που υπολογίζονται από τη μονάδα MULT τροφοδοτούνται σε μια μονάδα Εύρεσης της Μικρότερης Τιμής (Find Minimum Value). Η μονάδα Find Min δέχεται ανά κύκλο ρολογιού τους χρόνους t_j και στο τέλος τροφοδοτεί στην έξοδο της τη μικρότερη τιμή t_j δηλαδή τη τιμή t_μ . Η λειτουργία της μονάδας Find Min έχει ως εξής: αρχικά συγκρίνει τη πρώτη τιμή t_j που δέχεται με τη μεγαλύτερη επιτρεπτή Floating Point τιμή. Η τιμή αυτή είναι μια σταθερά η οποία έχει καθοριστεί πριν τη σύνθεση του συστήματος. Εάν η πρώτη τιμή που δέχεται είναι μικρότερη από τη προκαθορισμένη σταθερά τότε την αποθηκεύει σε έναν καταχωρητή αλλιώς περιμένει την επόμενη τιμή μέχρι να τροφοδοτηθεί κάποια που θα είναι μικρότερη από τη σταθερά. Έπειτα, σε κάθε κύκλο ρολογιού συγκρίνεται η τρέχουσα t_j τιμή με την t_{j-1} και η μικρότερη τιμή αποθηκεύεται στον καταχωρητή όπως προαναφέρθηκε. Συνεπώς, η μονάδα Find Min έχει πάντοτε αποθηκευμένη τη μικρότερη τιμή του χρόνου t_j και με την ολοκλήρωση των υπολογισμών όλων των t_j χρόνων τροφοδοτείται στην έξοδο της μονάδας η ελάχιστη τιμή του χρόνου t_j (t_μ).

Μια επιπλέον λειτουργία της μονάδας Find Min είναι να εντοπίσει το δείκτη μ της αντίδρασης R_μ από την οποία προέκυψε ο χρόνος t_μ . Για το λόγο αυτό χρησιμοποιεί έναν μετρητή που αυξάνει με κάθε νέα τιμή t_j ενώ αποθηκεύει τη τιμή του μετρητή μόνο εάν $t_j < t_{j-1}$. Στην Εικόνα 19 ο δείκτης μ της R_μ αντίδρασης συμβολίζεται ως WRA (Winning Reaction Address). Το εύρος του μετρητή εξαρτάται από το μέγεθος του τμήματος του πίνακα RT που διαχειρίζεται η εκάστοτε FPU.

Όπως αναφέρθηκε, η μονάδα PE είναι αρχιτεκτονική πλήρους διοχέτευσης η οποία μπορεί να επεξεργάζεται μια αντίδραση του πίνακα RT σε κάθε κύκλο ρολογιού. Μετά από τη τυπική καθυστέρηση λόγω των επιπέδων διασωλήνωσης, η PE μονάδα υπολογίζει μια νέα τιμή του χρόνου t_j σε κάθε κύκλο ρολογιού. Τα δύο μονοπάτια της διόδου δεδομένων ξεκινάν τη λειτουργίας τους ταυτόχρονα και επεξεργάζονται τα δεδομένα μιας αντίδρασης παράλληλα ενώ συγχρονίζονται μέσω σημάτων ελέγχου.

5.5 Propensity Calculation Unit - PCU

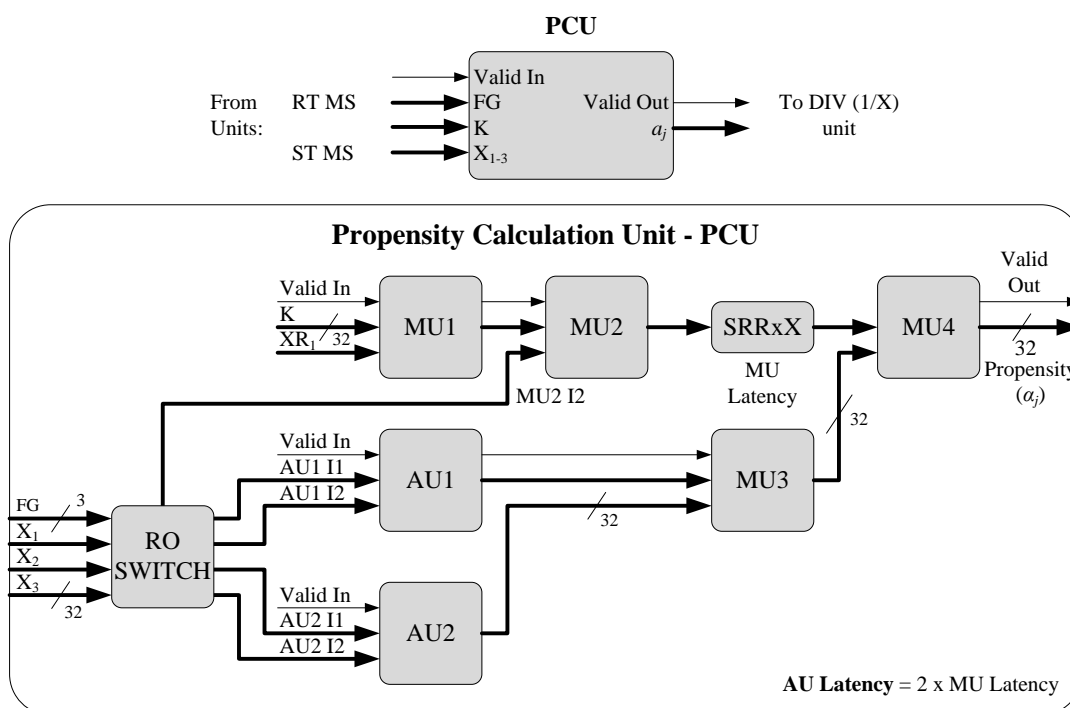
Η σχέση υπολογισμού του χρόνου t_j απαιτεί πρώτα τον υπολογισμό της τάσης (propensity) a_j μιας αντίδρασης R_j . Όπως αναφέρθηκε και Στην Ενότητα των 3, το σύστημα FRM SoC μπορεί να υπολογίζει τάσεις αντιδράσεων μέχρι και τρίτης τάξης (με τρία αντιδρώντα). Μελετώντας τις ODEs για κάθε πιθανό συνδυασμό αντιδράσεων έως τρίτης τάξης καταλήξαμε στις σχέσεις του Πίνακα 5 για τον υπολογισμό των τάσεων [57]:

Πίνακας 5: Υπολογισμός των τάσεων (propensities) των αντιδράσεων έως τρίτης τάξης

Order	Reaction	Propensity Calculation
1	$X_1 \xrightarrow{K} Y_i + \dots$	$K \cdot X_1$
2.a	$X_1 + X_1 \xrightarrow{K} Y_i + \dots$	$K \cdot X_1 \cdot (X_1 - 1) \cdot 0.5$
2.b	$X_1 + X_2 \xrightarrow{K} Y_i + \dots$	$K \cdot X_1 \cdot X_2$
3.a	$X_1 + X_2 + X_3 \xrightarrow{K} Y_i + \dots$	$K \cdot X_1 \cdot X_2 \cdot X_3$
3.b	$X_1 + X_1 + X_2 \xrightarrow{K} Y_i + \dots$	$K \cdot X_1 \cdot (X_1 - 1) \cdot 0.5 \cdot X_2$
3.c	$X_1 + X_1 + X_1 \xrightarrow{K} Y_i + \dots$	$K \cdot X_1 \cdot (X_1 - 1) \cdot (X_1 - 2) \cdot 0.16$

Όπου K είναι η σταθερά του ρυθμού αντίδρασης και $X_1 - X_3$ οι μοριακοί πληθυσμοί των αντιδρώντων μιας αντίδρασης.

Για τον υπολογισμό των τάσεων των αντιδράσεων σχεδιάστηκε μια μονάδα (Propensity Calculation Unit – PCU) επίσης πλήρους διοχέτευσης. Στη συνέχεια ακολουθεί σχηματικό αρχιτεκτονικής της μονάδας PCU.



Εικόνα 20: Αρχιτεκτονική της μονάδας Propensity Calculation Unit - PCU

Η μονάδα σχεδιάστηκε ώστε να μπορεί να τροφοδοτείται με νέα δεδομένα σε κάθε κύκλο ρολογιού. Μετά από τη τυπική καθυστέρηση των επιπέδων διασολήνωσης η μονάδα PCU θα αρχίσει να παράγει σε κάθε κύκλο ρολογιού μια νέα τάση αντίδρασης για όσο χρονικό διάστημα λαμβάνει στις εισόδους της έγκυρα δεδομένα. Η εγκυρότητα των δεδομένων σηματοδοτείται μέσω του σήματος *Valid In* το οποίο μεταδίδεται σταδιακά σε όλες

τις μονάδες της δίοδου δεδομένων προκειμένου να ενεργοποιηθούν οι μονάδες της για το χρονικό διάστημα που αυτό παραμένει στη λογική κατάσταση (1). Η συνδεσμολογία της Εικόνας 14 προέκυψε από τις παρακάτω προδιαγραφές σχετικά με τη λειτουργία της κάθε μονάδας της αρχιτεκτονικής της PCU.

- MU1: υπολογίζει πάντοτε το γινόμενο $K \times X_1 - MU1 \text{ Out}$
- MU2: υπολογίζει το γινόμενο $MU1 \text{ Out} \times 1$ ή $\times 0.5$ ή $\times 0.16$
- MU3: υπολογίζει το γινόμενο $AU1 \text{ Out} \times AU2 \text{ Out}$
- MU4: υπολογίζει το γινόμενο $MU2 \text{ Out} \times MU3 \text{ Out}$
- AU1: υπολογίζει το άθροισμα $(1 + 0)$ ή $(X_1 + (-1))$ ή $(X_2 + 0)$
- AU2: υπολογίζει το άθροισμα $(1 + 0)$ ή $(X_2 + 0)$ ή $(X_3 + 0)$ ή $(X_1 + (-2))$

Ο ολισθητής SRRxX της Εικόνας 20 απαιτείται προκειμένου να συγχρονίζονται τα δεδομένα του κάτω μονοπατιού με τα δεδομένα του πάνω μονοπατιού. Η χρήση του ολισθητή είναι απαραίτητη λαμβάνοντας υπόψη ότι οι αθροιστές έχουν διπλάσιο χρόνο καθυστέρησης, λόγω διασωλήνωσης, από τους πολλαπλασιαστές.

Η παραπάνω λειτουργία εξασφαλίζεται από την υπομονάδα Reaction Order Switch – RO SWITCH η οποία είναι υπεύθυνη για την ορθή τροφοδοσία δεδομένων των δύο εισόδων των αθροιστών AU1, AU2 και της μιας εισόδου του πολλαπλασιαστή MU2. Η μονάδα RO SWITCH δέχεται τους μοριακούς πληθυσμούς των τριών αντιδρώντων καθώς και το σήμα FG το οποίο κωδικοποιεί τις περιπτώσεις του πίνακα των τάσεων των αντιδράσεων. Στον Πίνακα 6 που ακολουθεί δίνονται οι τιμές με τις οποίες τροφοδοτεί η μονάδα RO SWITCH τις μονάδες AU1, AU2 και MU2 προκειμένου να υπολογιστούν οι τάσεις των αντιδράσεων σύμφωνα με το πίνακα:

Πίνακας 6: Λειτουργίες της υπομονάδας Reaction Order SWITCH της PCU

Reaction Order SWITCH						
	MU1		AU1		AU2	
FG	In 1	In 2	In 1	In 2	In 1	In 2
1	MU1	· 1	1	+ 0	1	+ 0
2.a	MU1	· 0.5	X_1	+ (-1)	1	+ 0
2.b	MU1	· 1	X_2	+ 0	1	+ 0
3.a	MU1	· 1	X_2	+ 0	X_3	+ 0
3.b	MU1	· 0.5	X_1	+ (-1)	X_2	+ 0
3.c	MU1	· 0.16	X_1	+ (-1)	X_1	+ (-2)

Στην μονάδα PCU μπορούν να διοχετεύονται νέες αντιδράσεις διαφορετικών τάξεων σε κάθε κύκλο ρολογιού. Αυτό επιτυγχάνεται από την αρχιτεκτονική της μονάδας PCU και τα μονοπάτια δεδομένων (dataflow) της. Και τα δύο μονοπάτια δεδομένων της αρχιτεκτονικής (επάνω και κάτω) έχουν τον ίδιο συνολικά αριθμό καταχωρητών, ανεξαρτήτου της τάξεως μιας αντίδρασης, εξασφαλίζοντας σταθερή καθυστέρηση λόγω διοχέτευσης (latency) στην διάδοση των δεδομένων. Συνεπώς η διοχέτευση των αντιδράσεων στις υπομονάδες της PCU δεν εξαρτάται από την τάξη της εκάστοτε αντίδρασης. Έστω για παράδειγμα ότι διοχετεύεται μια

αντίδραση τρίτης τάξης, τριών διαφορετικών αντιδρώντων. Ο υπολογισμός της τάσης της, σύμφωνα και με τον Πίνακα 5, δεν απαιτεί την πράξη της άθροισης. Παρόλα αυτά οι πληθυσμοί X_2 και X_3 θα διοχετευθούν στους αθροιστές AU1 και AU2, όπως φαίνεται και στον Πίνακα 6 για τη περίπτωση 3.α, ωστόσο δεν θα μεταβληθούν διότι η μονάδα Reaction Order SWITCH θα εξασφαλίσει στη περίπτωση αυτή να γίνει άθροιση με το μηδέν. Η σχεδίαση αυτή επιτυγχάνει να διοχετεύονται νέες αντιδράσεις, διαφορετικών τάξεων, σε κάθε κύκλο ρολογιού ενώ μετά από την αρχική καθυστέρηση λόγω διοχέτευσης (latency) θα παράγονται τάσεις α_j επίσης σε κάθε κύκλο ρολογιού.

5.6 Θεωρητική Ανάλυση Επιδόσεων

Στις προηγούμενες ενότητες έγινε εκτενής παρουσίαση των τεχνικών περιγραφών όλων των κύριων μονάδων του FRM πυρήνα και του FRM SoC. Στην ενότητα που ακολουθεί παρουσιάζονται τα αποτελέσματα της θεωρητικής ανάλυσης της καθυστέρησης λόγω διοχέτευσης (latency) της αρχιτεκτονικής και το πώς αυτή επηρεάζει την συνολική απόδοση του συστήματος.

Ανεξαρτήτως από τη καθυστέρηση της κάθε μονάδας θα πρέπει να εξεταστεί τι καθορίζει την συνολική απόδοση του συστήματος FRM SoC. Ένα βιολογικό σύστημα μελετάται από τις καταστάσεις μετάβασης του. Οι καταστάσεις μετάβασης του συστήματος σηματοδοτούνται από τα γεγονότα (reaction events) εκείνα που προκαλούν αυξομειώσεις στους μοριακούς πληθυσμούς των μοριακών ειδών και αύξηση του χρόνου προσομοίωσης t_{sim} . Τέτοια γεγονότα στον FRM-SSA αλγόριθμο προκύπτουν μετά το τέλος ενός κύκλου προσομοίωσης (RC) δηλαδή της επεξεργασίας ολόκληρου του συνόλου των αντιδράσεων του πίνακα RT. Στην Ενότητα 2.4, ορίστηκε ο κύκλος αντιδράσεων (reaction cycle) κατά τον οποίο εκτελούνται τα βήματα 2 έως 5 του FRM-SSA αλγόριθμου όπου και εξετάζονται όλες οι αντιδράσεις του βιομοντέλου προκειμένου να υπολογιστεί η επόμενη κατάσταση του βιολογικού συστήματος. Εφόσον η εξέλιξη του βιολογικού συστήματος καθορίζεται από την ολοκλήρωση κύκλων αντιδράσεων χρησιμοποιείται το παρακάτω μέγεθος για τη θεωρητική μελέτη της διεκπεραιωτικής ικανότητας (throughput) του FRM SoC συστήματος.

- **Throughput: Mega Reactions Cycles / sec**, το μέγεθος **MRC/sec** καθορίζει το πόσα εκατομμύρια κύκλοι αντιδράσεων (RCs) ολοκληρώνονται σε ένα δευτερόλεπτο στο σύστημα FRM SoC.

Το μέγεθος MRC/sec εξαρτάται άμεσα από το πλήθος των αντιδράσεων ενός βιομοντέλου αφού η ολοκλήρωση ενός κύκλου αντιδράσεων προϋποθέτει την επεξεργασία όλου του συνόλου των αντιδράσεων (m) του πίνακα RT. Προκειμένου να μελετήσουμε την επίδοση (performance) του FRM SoC συστήματος

ανεξάρτητα από το πλήθος των αντιδράσεων ενός βιομοντέλου χρησιμοποιούμε και ένα δεύτερο μέγεθος

- **Performance: Mega Reactions / sec**, το μέγεθος **MR/sec** καθορίζει το πόσα εκατομμύρια αντιδράσεις εξετάζονται σε ένα δευτερόλεπτο από το σύστημα FRM SoC.

Το μέγεθος MR/sec προκύπτει από το γινόμενο του μεγέθους MRC/sec με τον συνολικό αριθμό αντιδράσεων του μοντέλου (m). Για την εκτίμηση του μεγέθους MRC/sec θα πρέπει εξεταστεί το πόσοι κύκλοι ρολογιού απαιτούνται για την ολοκλήρωση ενός κύκλου αντιδράσεων σε συνάρτηση με τη συχνότητα λειτουργίας του FRM SoC συστήματος. Το μέγεθος που αφορά τους κύκλους ρολογιού που απαιτούνται για την ολοκλήρωση ενός κύκλου αντιδράσεων το ορίζουμε ως RC Latency – L_{RC} . Για την ανάλυση του L_{RC} αρχικά θα πρέπει να

εξεταστούν οι καθυστερήσεις λόγω διοχέτευσης των μονάδων εκείνων που σχετίζεται με την εκτέλεση ενός κύκλου αντιδράσεων. Στη συνέχεια ακολουθεί πίνακας με τις καθυστερήσεις των μονάδων αυτών.

Πίνακας 7: Καθυστερήσεις λόγω διοχέτευσης της PE μονάδας

PE Datapath Level - Unit Latency					
Level	PE Datapath - Upper Branch				
Unit	MT19937	F2FLT	LNU	Total	
Latency (cc)	2	1	23	26	
Level	PE Datapath - Lower Branch				
Unit	PCU	DIV 1/X	MULT	FIND MIN	Total
Latency (cc)	16	26	4	1	47

Πίνακας 8: Καθυστερήσεις λόγω διοχέτευσης σε επίπεδο FRM πυρήνα

FRM CORE / FMU / FPU Level - Unit Latency	
Level - Unit	Latency (cc)
FPU – RT MS	2
FPU – ST MS	2
FMU - COMP TREE	$\log_2(N)$
FRM CORE - STU	3

Στους παραπάνω πίνακες έχουν συμπεριληφθεί όλες οι καθυστερήσεις των υπομονάδων που απαρτίζουν την FMU μονάδα συν τη καθυστέρηση της μονάδας STU η οποία είναι υπεύθυνη για την ανανέωση των μοριακών πληθυσμών και του χρόνου προσομοίωσης t_{sim} . Οι υπόλοιπες μονάδες του FRM πυρήνα δεν συμβάλουν άμεσα στην αύξηση της καθυστέρησης και για αυτό δε περιλαμβάνονται οι καθυστερήσεις τους. Στο σημείο αυτό θα πρέπει να γίνει ένας σημαντικός διαχωρισμός. Οι κύκλοι ρολογιού των Πινάκων 6 και 7 αντιπροσωπεύουν τους κύκλους ρολογιού που απαιτούνται έως ότου οδηγηθούν τα πρώτα έγκυρα δεδομένα στις εξόδους των παραπάνω μονάδων. Ωστόσο για τον υπολογισμό του L_{RC} θα πρέπει να συμπεριληφθούν και οι κύκλοι ρολογιού που απαιτούνται για την ολοκλήρωση της λειτουργίας των μονάδων που σχετίζονται με αυτή τη καθυστέρηση. Επίσης, όπως φαίνεται και στο Πίνακα 5 το πάνω μονοπάτι (upper branch) της μονάδας PE Datapath δεν επηρεάζει το L_{RC} διότι η καθυστέρηση του είναι πάντοτε μικρότερη σε σύγκριση με το κάτω μονοπάτι (lower branch). Για το λόγο αυτό δίνεται και ένας τρίτος πίνακας στον οποίο έχουν συμπεριληφθεί μόνο τα μονοπάτια των δεδομένων (dataflow) τα οποία καθορίζουν το L_{RC} . Επίσης σε αυτόν τον πίνακα συμπεριλαμβάνονται και οι κύκλοι ρολογιού που απαιτούνται για την ολοκλήρωση των λειτουργιών των μονάδων που επηρεάζουν το L_{RC} .

Πίνακας 9: Καθυστερήσεις λόγω διοχέτευσης της διόδου δεδομένων

Reaction Cycle Latency					
Level/Unit	FPU – PE Datapath		FMU	FRM CORE	
Dataflow	Data Fetch	Lower Branch	CMP TREE	STU	Others
Latency (cc)	5	$47+(m/N)$	$\text{Log}_2(N)$	25	6

Η δεύτερη στήλη (Data Fetch) του παραπάνω πίνακα δείχνει τους κύκλους ρολογιού που απαιτούνται για τη προσκόμιση των δεδομένων στη μονάδα PE. Οι 5 κύκλοι ρολογιού προκύπτουν από τη καθυστέρηση της μονάδας RT MS συν τη καθυστέρηση της μονάδας ST MS συν ένα επιπλέον κύκλο λόγω ενδιάμεσου καταχωρητή. Οι 25 κύκλοι ρολογιού στη στήλη FRM CORE – STU προκύπτουν από τους κύκλους ρολογιού που απαιτούνται για την ολοκλήρωση της λειτουργίας της μονάδας STU. Οι έξι κύκλοι ρολογιού στη τελευταία στήλη του Πίνακα 7 είναι καθυστερήσεις που οφείλονται σε επιπλέον καταχωρητές της διόδου δεδομένων και σε καθυστερήσεις μονάδων ελέγχου του FRM πυρήνα που λειτουργούν με σύγχρονο τρόπο. Οι κύκλοι ρολογιού για την ολοκλήρωση του κάτω μονοπατιού (lower branch) της μονάδας PE εξαρτώνται από τη σταθερή καθυστέρηση 47 κύκλων ρολογιού λόγω διοχέτευσης συν τους m/N κύκλους ρολογιού που απαιτούνται για την επεξεργασία των αντιδράσεων του πίνακα RT. Η καθυστέρηση του δέντρου συγκριτών εξαρτάται από το βάθος του δέντρου ($\text{Log}_2(N)$) εφόσον μεσολαβούν καταχωρητές σε κάθε επίπεδο του δέντρου.

Είναι εμφανές από όλα τα παραπάνω ότι η ολοκλήρωση ενός κύκλου αντιδράσεων εξαρτάται από μια σταθερή καθυστέρηση λόγω διοχέτευσης των μονάδων που επεξεργάζονται το σύνολο των αντιδράσεων του πίνακα RT συν τους κύκλους ρολογιού που απαιτούνται για τη προσκόμιση τους. Η FPU μονάδα είναι σχεδιασμένη έτσι ώστε να μπορεί να εξετάζει μια νέα αντίδραση R_j σε κάθε κύκλο ρολογιού. Αυτό έχει ως αποτέλεσμα, μετά από την αρχική καθυστέρηση λόγω διοχέτευσης να υπολογίζονται νέοι χρόνοι τ_j από τη μονάδα FPU σε κάθε κύκλο ρολογιού. Συνεπώς, αθροίζοντας όλες τις σταθερές καθυστερήσεις (K) καταλήγουμε στη παρακάτω σχέση για τον υπολογισμό των κύκλων ρολογιού που απαιτούνται για την ολοκλήρωση ενός κύκλου αντιδράσεων.

(3)

$$L_{RC} = \frac{m}{N} + \text{Log}_2(N) + K$$

όπου $K = 83$

Το αποτέλεσμα της παραπάνω σχέσης μας δίνει το συνολικό αριθμό κύκλων ρολογιού που μεσολαβούν από την έναρξη του τρέχοντος κύκλου αντιδράσεων έως την έναρξη του επόμενου κύκλου αντιδράσεων.

Χρησιμοποιώντας τη Σχέση 4 μπορούμε να μελετήσουμε θεωρητικά τις επιδόσεις του FRM SoC καθώς η αρχιτεκτονική κλιμακώνεται σε επίπεδο επεξεργαστικών μονάδων FPUs συναρτήσεως του πλήθους των αντιδράσεων δεδομένου μιας σταθερής συχνότητας. Η διεκπεραιωτική ικανότητα (Throughput-MRC/sec) του FRM SoC συστήματος μπορεί να υπολογιστεί μέσω του παρακάτω τύπου:

(4)

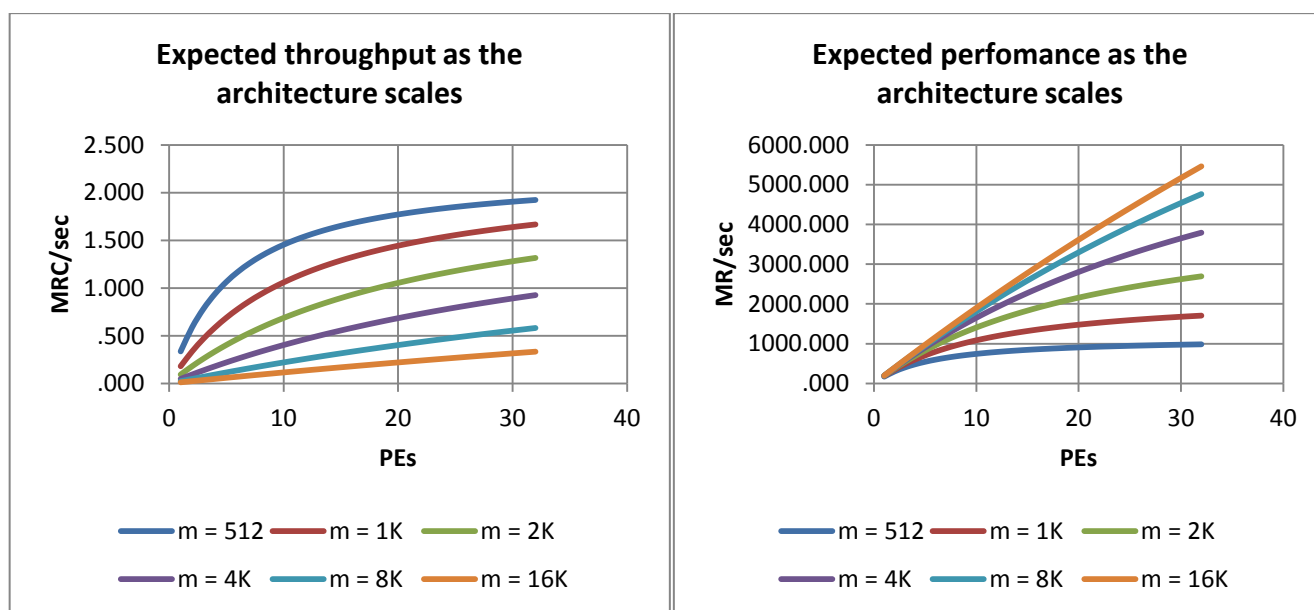
$$T = \frac{1}{RC_{Lat} \cdot t_{cycle} \cdot 10^6} (MRC/sec)$$

Όπως έχει ήδη αναφερθεί ο υπολογισμός της απόδοσης (Performance-MR/sec) προκύπτει από το γινόμενο της ρυθμαπόδοσης με το πλήθος των αντιδράσεων.

(6)

$$P = T \cdot m (MR/sec)$$

Στη συνέχεια ακολουθούν δύο διαγράμματα για την θεωρητική επίτευξη της διεκπεραιωτικής ικανότητας (throughput) και της επίδοσης (performance) του FRM SoC συστήματος για διάφορους συνδυασμούς αριθμών αντιδράσεων του βιομοντέλου (m) και διαθέσιμων επεξεργαστικών μονάδων (N).



Εικόνα 21: Θεωρητικά αναμενόμενη διεκπεραιωτική ικανότητα και επίδοση του συστήματος FRM SoC ως συνάρτηση του πλήθους αντιδράσεων του βιομοντέλου (m) και του αριθμού των PEs (N). Συχνότητα λειτουργίας 175 MHz

Για τον υπολογισμό των διαγραμμάτων της Εικόνα 21 έχει χρησιμοποιηθεί η σταθερή συχνότητας λειτουργίας των 175 MHz. Όπως θα αναφερθεί και στην Ενότητα 6, όπου παρουσιάζονται τα στατιστικά αποτελέσματα της αρχιτεκτονικής, η συχνότητα των 175 MHz είναι μια εφικτή συχνότητα λειτουργίας του συστήματος FRM SoC στο FPGA που χρησιμοποιήθηκε κατά την ανάπτυξη του. Ο οριζόντιος άξονας αντιπροσωπεύει την κλιμάκωση των επεξεργαστικών μονάδων FPU. Ο κάθετος άξονας στο αριστερό γράφημα αντιπροσωπεύει τη διεκπεραιωτική ικανότητα (throughput) και ο κάθετος άξονας στο δεξιό γράφημα αντιπροσωπεύει την επίδοση (performance) του FRM SoC συστήματος. Επιπλέον, η ανάλυση αυτή έγινε για μια ποικιλία πλήθους αντιδράσεων της μορφής $m = 2^k$ για $k = 9$ έως 14. Παρατηρείται ότι η απόδοση του συστήματος για $N = 32$

τείνει να αυξάνεται γραμμικά όσο το μέγεθος του μοντέλου μεγαλώνει, επιβεβαιώνοντας τη χρησιμότητα του FRM SoC για μεγάλο μέγεθος προσομοιώσεις με χιλιάδες αντιδράσεις.



6. ΣΤΑΤΙΣΤΙΚΑ ΥΛΟΠΟΙΗΣΗΣ

Στις δύο προηγούμενες ενότητες παρουσιάστηκε εκτενώς η αρχιτεκτονική FRM SoC, σε αυτήν την ενότητα παρουσιάζονται τα στατιστικά υλοποίησης του FRM SoC για διάφορες τιμές των VHDL Generic παραμέτρων. Για τη σύνθεση και την υλοποίηση του συστήματος χρησιμοποιήθηκε το λογισμικό ISE 14.5 [44] της εταιρίας Xilinx [38]. Όλα τα στατιστικά που παρουσιάζονται σε αυτήν την ενότητα έχουν προκύψει από το στάδιο της υλοποίησης (Implementation) του λογισμικού ISE 14.5. Η υλοποίηση του συστήματος έγινε με το FPGA XC7K325T-2FFG900 της οικογένειας Kintex7 της εταιρίας Xilinx. Στη συνέχεια ακολουθεί ένας πίνακας με τα κύρια χαρακτηριστικά του FPGA XC7K325T.

Πίνακας 10: Χαρακτηριστικά του FPGA XC7K325T

XC7K325T				
Flip-Flops	LUTs	BRAMs (36Kbit)	DSPs	IOBs
407,600	203,800	445	840	500

Στη συνέχεια ακολουθεί ένας πίνακας με στατιστικά υλοποίησης του FRM πυρήνα και των υπομονάδων του (Πίνακας 11), ένας πίνακας με στατιστικά υλοποίησης του PCIe πυρήνα και της μονάδας AXIS4 SWITCH (Πίνακας 12) και δύο πίνακες με στατιστικά υλοποίησης για τους δύο τρόπους λειτουργίας SSIP (Πίνακας 13) και MSIP (Πίνακας 14). Όλα τα στατιστικά υλοποίησης που ακολουθούν έχουν προκύψει για τις μέγιστες τιμές των παραμέτρων m και n δεδομένου των 445 BRAM των 36-Kbit που διαθέτει το FPGA XC7K325T.

Πίνακας 11: Στατιστικά υλοποίησης του FRM πυρήνα

Στατιστικά υλοποίησης του FRM πυρήνα ($m = n = 2^{12}$)					
	Flip Flops	LUTs	BRAMs	DSPs	Freq. MHz
Core Gen Units					
F2FLT	217 (1%)	177 (1%)	–	–	484
FPADDER	202 (1%)	428 (1%)	–	2 (1%)	300
FPMULT	72 (1%)	89 (1%)	–	3 (1%)	367
FPDIV	688 (1%)	858 (1%)	–	–	301
FPLN	1,114 (1%)	925 (1%)	–	4 (1%)	442
PE Level					
PCU	772 (1%)	917 (1%)	–	16 (1%)	270
MT19937	46 (1%)	196 (1%)	2 (1%)	–	258
Find Min	82 (1%)	131 (1%)	–	–	285
PE	3,091 (1%)	3,247 (1%)	2 (1%)	23 (2%)	258
FPU Level					
RT MS	229 (1%)	246 (1%)	22 (4%)	–	260
ST MS	54 (1%)	103 (1%)	10 (2%)	–	265
FPCU	105 (1%)	129 (1%)	–	–	323
FPU	3,461 (1%)	3,687 (1%)	29 (6%)	23 (2%)	283
Core Level					

STI	38 (1%)	58 (1%)	4 (1%)	–	260
FMU	3,537 (1%)	3,752 (1%)	29 (6%)	23 (2%)	257
SPU	716 (1%)	859 (1%)	4 (1%)	2 (1%)	280
SCU	338 (1%)	412 (1%)	–	–	300
IIS	307 (1%)	368 (1%)	–	–	310
OIS	684 (1%)	650 (1%)	8 (1%)	–	280
FRM Core	5,541 (1%)	5,828 (2%)	44 (9%)	25 (2%)	255

Πίνακας 12: Στατιστικά υλοποίησης του PCIe πυρήνα και της μονάδας AXIS4 SWITCH

Στατιστικά υλοποίησης του PCIe πυρήνα και της μονάδας AXIS4 SWITCH					
	Flip Flops	LUTs	BRAMs	DSPs	Freq. MHz
PCIe CORE	7,648 (1%)	8,078 (1%)	12 (2%)	–	200
AXIS4 SWITCH	106 (1%)	335 (1%)	–	–	330

Πίνακας 13: Στατιστικά υλοποίησης του FRM1xN SoC συστήματος - SSIP

Στατιστικά υλοποίησης του FRM1xN SoC συστήματος - SSIP ($m = n = 2^{12}$)					
FRM NoC 1xN	Flip Flops	LUTs	BRAMs	DSPs	Freq. MHz
1x1	12,705 (3%)	12,700 (6%)	23 (5%)	25 (2%)	255
1x2	18,591 (4%)	21,808 (10%)	68 (15%)	48 (5%)	250
1x4	25,795 (6%)	30,586 (15%)	89 (20%)	94 (11%)	250
1x8	40,174 (9%)	48,051 (23%)	149 (33%)	186 (22%)	200
1x16	68,896 (16%)	82,355 (40%)	229 (51%)	370 (44%)	190
1x32	124,281 (30%)	128,441 (63%)	397 (87%)	738 (87%)	175

Πίνακας 14: Στατιστικά υλοποίησης του FRMCx1 SoC συστήματος - MSIP

Στατιστικά υλοποίησης του FRMCx1 SoC συστήματος - MSIP ($m = n = 2^{11}$)					
FRM NoC Cx1	Flip Flops	LUTs	BRAMs	DSPs	Freq. MHz
1x1	12,705 (3%)	12,700 (6%)	35 (7%)	25 (2%)	255

2x1	18,063 (4%)	18,306 (8%)	46 (10%)	50 (5%)	255
4x1	28,777 (7%)	29,327 (14%)	74 (17%)	100 (11%)	253
8x1	50,306 (12%)	51,509 (25%)	135 (30%)	200 (23%)	253
16x1	93,357 (22%)	95,473 (46%)	236 (53%)	400 (47%)	220

Ο Πίνακας 11 δείχνει ότι FRM πυρήνας απαιτεί το 2% των διαθέσιμων LUTs και το 1% των διαθέσιμων Flip-Flops ενώ μπορεί να λειτουργήσει στη συχνότητα των 255 MHz για το συγκεκριμένο FPGA. Το σύστημα επικοινωνίας (PCIe core) και ο διανομέας του δικτύου έχουν επίσης ελάχιστες απαιτήσεις σε πόρους υλικού. Ο PCIe πυρήνας είναι χρονισμένος στα 200 MHz και η συχνότητα αυτή δεν αποτελεί άνω όριο για το FRM SoC διότι χρησιμοποιεί διαφορετικό σήμα ρολογιού. Ο μέγιστος αριθμός επεξεργαστικών μονάδων FPU (N) που πορούν να υλοποιηθούν στο συγκεκριμένο FPGA είναι οι 32 (Πίνακας 13). Ο μέγιστος αριθμός FRM πυρήνων που μπορούν υλοποιηθούν στο FPGA είναι οι 16 και περιορίζονται αποκλειστικά από τις διαθέσιμες BRAMs (Πίνακας 14).

Από τον Πίνακα 13 προκύπτει ότι τρόπος SSIP για $N = 32$ επιτυγχάνει μέγιστη συχνότητα λειτουργίας 175 MHz ενώ ο τρόπος MSIP επιτυγχάνει καλύτερη συχνότητα λειτουργίας στα 220 MHz για $C = 16$ (Πίνακας 14). Μια σημαντική διαφοροποίηση των δύο τρόπων είναι το δέντρο συγκριτών (CMP TREE) της μονάδας FMU, που απουσιάζει από τον τρόπο MSIP, καθώς σε αυτό εμφανίζονται μεγάλα συνδυαστικά μονοπάτια σύμφωνα με την χρονική ανάλυση (timing report) του λογισμικού ISE 14.5.

7. ΕΠΙΚΥΡΩΣΗ ΣΧΕΔΙΑΣΗΣ

Στις προηγούμενες ενότητες παρουσιάστηκε η αρχιτεκτονική FRM SoC και τα στατιστικά υλοποίησης για διαφορετικούς συνδυασμούς των τρόπων λειτουργίας (SSIP, MSIP). Στη παρούσα ενότητα θα παρουσιαστεί η μεθοδολογία που ακολουθήθηκε για τον έλεγχο και την επικύρωση της λειτουργίας της αρχιτεκτονικής. Επίσης θα παρουσιαστούν τα αποτελέσματα των δοκιμών που διεξήχθησαν χρησιμοποιώντας βιομοντέλα τύπου θηράματος/θηρευτή. (Lotka Volterra System – LVS [52]).

Κατά το στάδιο της ανάπτυξης κάθε μονάδα που υλοποιούταν ελεγχόταν (verification [58]) προκειμένου να επαληθευθεί ότι πληροί τις σχεδιαστικές προδιαγραφές. Για το λόγο αυτό χρησιμοποιήθηκε το λογισμικό προσομοίωσης Xilinx ISim [59] και κατάλληλα προγράμματα επαλήθευσης (VHDL testbenches). Τα προγράμματα αυτά τροφοδοτούν με διανύσματα δεδομένων (test vectors) τις μονάδες προς επαλήθευση. Έπειτα, μέσω της προσομοίωσης, ελέγχονται οι κυματομορφές των σημάτων των μονάδων προκειμένου να επαληθευθεί εάν τηρούν τις σχεδιαστικές προδιαγραφές. Το στάδιο αυτό χωρίζεται σε δύο φάσεις. Σε πρώτη φάση ελέγχεται η λογική συμπεριφορά των μονάδων μέσω προσομοιώσεων συμπεριφοράς (behavioral simulations) [60]. Στη συνέχεια ελέγχονται οι χρονικοί περιορισμοί των κυκλωμάτων μέσω προσομοιώσεων χρονισμού σε επίπεδο πύλης (timing simulation) [61]. Ο κύκλος ανάπτυξης και επαλήθευσης τερματίζει όταν μια μονάδα φτάσει στο σημείο να πληροί όλες τις σχεδιαστικές προδιαγραφές. Έπειτα και αφού πρώτα έχει επαληθευθεί η συμπεριφορά της κάθε μονάδας μπορεί να προστεθεί στο συνολικό σύστημα. Τέλος δοκιμάστηκε και επαληθεύθηκε συνολικά η λειτουργία του FRM SoC συστήματος και σε επίπεδο συμπεριφοράς αλλά και σε επίπεδο χρονισμού.

Για την επικύρωση (validation [58]) της αρχιτεκτονικής FRM SoC πραγματοποιήθηκαν δοκιμές με βιομοντέλα που έχουν γνωστή και αναμενόμενη συμπεριφορά. Επιπλέον, τα αποτελέσματα αυτών των δοκιμών τα συγκρίναμε με τα αντίστοιχα αποτελέσματα που παράγει ο δημοφιλής προσομοιωτής Copasi [53] για τα ίδια βιομοντέλα με τις ίδιες αρχικές συνθήκες. Στην συνέχεια δίνονται οι αρχικές συνθήκες του βιομοντέλου LVS που χρησιμοποιήθηκε για την επικύρωση της αρχιτεκτονικής:

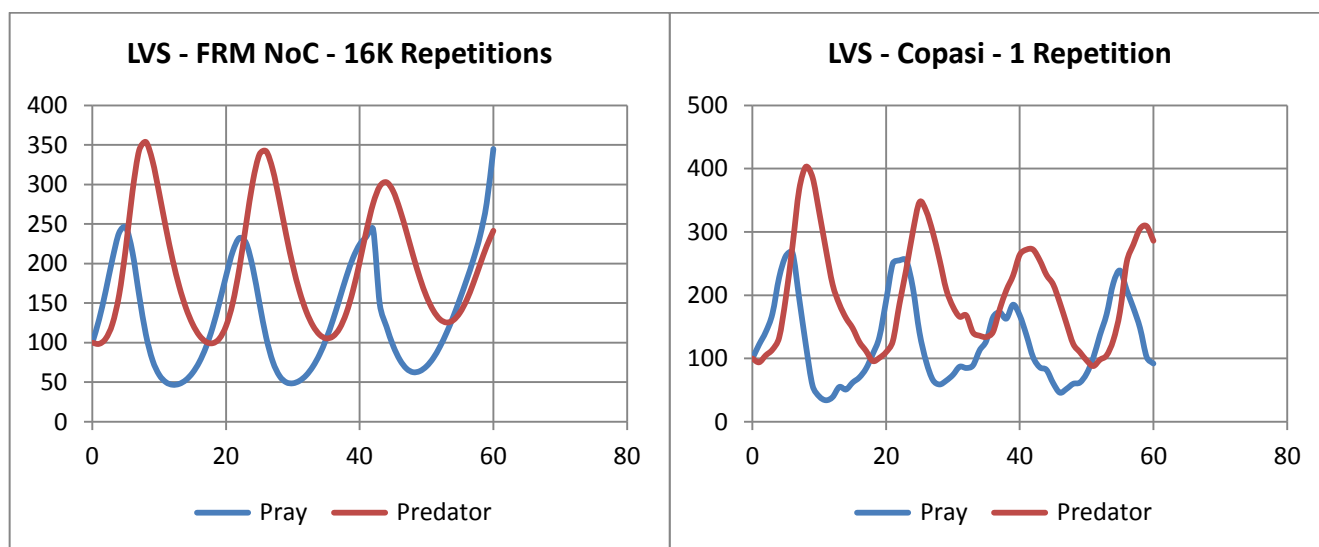
- Αρχικοί πληθυσμοί μοριακών ειδών:
 - Θηρευτής (Predator): 100
 - Θήραμα (Prey): 100
- Σταθερές ρυθμού αντίδρασης (K) για τις τρεις αντιδράσεις του βιομοντέλου LVS:
 - K_1 : 0,5
 - K_2 : 0,0025
 - K_3 : 0,3

Στη συνέχεια ακολουθούν οι παράμετροι συστήματος FRM SoC και οι παράμετροι προσομοίωσης που χρησιμοποιήθηκαν κατά τις δοκιμές με το βιομοντέλο LVS:

- **FRM NoC setup:** MSIP – 16x1
- **FRM NoC freq.:** 220 MHz
- **T_{sim} :** 60 s (διάρκεια προσομοίωσης)
- **T_{sam} :** 1 s (περίοδος δειγματοληψίας αποτελεσμάτων)

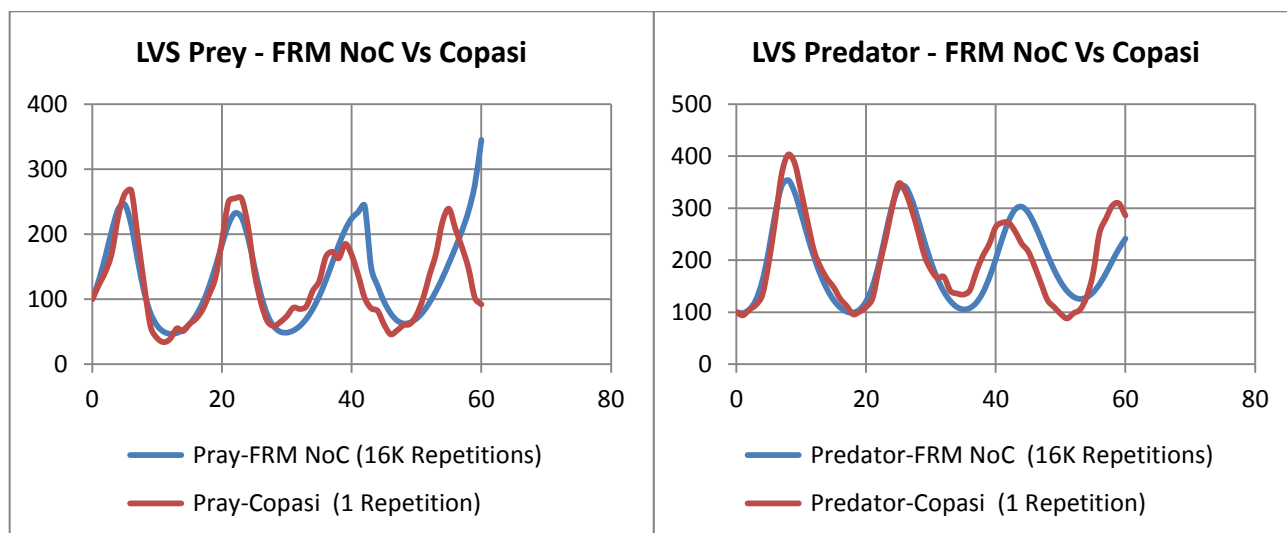
- **Repetitions:** R=16K (στοχαστικές επαναλήψεις)
- **Max Internal Steps:** 1000K

Για τη παραλληλοποίηση των 16K επαναλήψεων χρησιμοποιήθηκαν 16 FRM πυρήνες όπου ο καθένας είναι υπεύθυνος για την εκτέλεση 1000 επαναλήψεων. Κάθε σετ 1000 επαναλήψεων διαφοροποιείται πλήρως από τα υπόλοιπα εφόσον και οι 16 γεννήτριες ψευδοτυχαίων αριθμών των πυρήνων έχουν αρχικοποιηθεί με διαφορετικά διανύσματα αρχικοποίησης. Στη συνέχεια ακολουθούν τα αποτελέσματα των προσομοιώσεων που εκτελέστηκαν με το βιομοντέλο LVS στο σύστημα FRM NoC καθώς και στο προσομοιωτή Copasi.



Εικόνα 22: Αποτελέσματα προσομοιώσεων με το βιομοντέλο LVS

Οι γραφικές του FRM NoC συστήματος έχουν προκύψει από τη μέση τιμή των 16K επαναλήψεων. Η σύγκριση των αποτελεσμάτων των δύο προσομοιωτών φαίνεται καλύτερα στα δύο επόμενα γραφήματα.



Εικόνα 23: Σύγκριση των προσομοιωτών FRM NoC και Copasi

Στο αριστερό μέρος της Εικόνα 22 δίνονται τα αποτελέσματα προσομοίωσης του μοντέλου LVS για 16K επαναλήψεις στο FRM SoC σύστημα και στο δεξιό μέρος τα αποτελέσματα προσομοίωσης του ίδιου μοντέλου για μια επανάληψη στο προσομοιωτή Copasi. Στο αριστερό μέρος της Εικόνα 23 δίνονται οι μεταβολές του θηράματος (Prey) και στους δύο προσομοιωτές και αντίστοιχα στο δεξιό μέρος της εικόνας δίνονται οι μεταβολές του θηρευτή. Οι γραφικές παραστάσεις του FRM NoC εμφανίζονται να έχουν πιο ομαλές μεταβολές σε σύγκριση με τις γραφικές παραστάσεις του προσομοιωτή Copasi. Αυτό εξηγείται λόγω των χιλιάδων επαναλήψεων που εκτελέστηκε το πείραμα στον FRM NoC προσομοιωτή σε σύγκριση με τη μια επανάληψη που εκτελέστηκε στο προσομοιωτή Copasi. Παρόλα αυτά οι ομοιότητες στα αποτελέσματα των δύο προσομοιωτών είναι εμφανείς κάτι το οποίο μας οδηγεί στο συμπέρασμα ότι τουλάχιστον για το συγκεκριμένο σετ δοκιμών το FRM SoC σύστημα έχει την αναμενόμενη, ορθή συμπεριφορά και λειτουργία.

Στο FRM SoC η εκτέλεση $R=16K$ στοχαστικών επαναλήψεων διήρκησε $T_{real} = 12,3 \text{ sec}$ πραγματικού χρόνου (real time) ενώ ο προσομοιωτής Copasi χρειάστηκε 509 sec για να ολοκληρώσει το αντίστοιχο πείραμα. Ο προσομοιωτής Copasi εκτέλεσε τις 16K προσομοιώσεις σε επεξεργαστή τελευταίας γενιάς (Intel i7) με συχνότητα λειτουργίας 2,4 GHz ενώ είχε στη διάθεση του 8GB μνήμης RAM. Στη δοκιμή με το βιομοντέλο LVS το σύστημα FRM SoC με 16 πυρήνες στα 220 MHz αποδεικνύεται 41 φορές ταχύτερο (speedup factor) σε σύγκριση με τον προσομοιωτή Copasi.

Στη συνέχεια του έργου έχουμε προγραμματίσει επιπλέον δοκιμές με δημοσιευμένα μοντέλα από την Biomodels database. Επίσης θα διεξαχθούν δοκιμές και με μοντέλα benchmark, με ελεγχόμενη και κλιμακούμενη πολυπλοκότητα, με σκοπό την μελέτη των επιδόσεων που επιτυγχάνονται στη πράξη από την αρχιτεκτονική FRM SoC και την σύγκριση των αποτελεσμάτων αυτών με την θεωρητική εκτίμηση επιδόσεων που περιγράφεται στην Ενότητα 5.6.

ΑΝΑΦΟΡΕΣ

- [1] H. Kitano, «Systems biology: a brief overview,» *Science*, τόμ. 295, 2002.
- [2] Wikipedia contributors, «Systems Biology,» Wikipedia, the free encyclopedia. Wikimedia Foundation, Inc. [Ηλεκτρονικό]. [Πρόσβαση 09 Jun 2012].
- [3] Wikipedia contributors, «Computational biology,» Wikipedia, the free encyclopedia. Wikimedia Foundation, Inc. [Ηλεκτρονικό]. [Πρόσβαση 07 Jun 2012].
- [4] Wikipedia contributors, «Bioinformatics,» Wikipedia, the free encyclopedia. Wikimedia Foundation, Inc. [Ηλεκτρονικό]. [Πρόσβαση 07 Jun 2012].
- [5] Wikipedia contributors, «Data mining,» Wikipedia, the free encyclopedia. Wikimedia Foundation, Inc. [Ηλεκτρονικό]. [Πρόσβαση 08 Jun 2012].
- [6] Wikipedia contributors, «Control theory,» Wikipedia, the free encyclopedia. Wikimedia Foundation, Inc. [Ηλεκτρονικό]. [Πρόσβαση 29 May 2012].
- [7] Wikipedia contributors, «Network analysis (electrical circuits),» Wikipedia, the free encyclopedia. Wikimedia Foundation, Inc. [Ηλεκτρονικό]. [Πρόσβαση 29 May 2012].
- [8] Wikipedia contributors, «Artificial intelligence,» Wikipedia, the free encyclopedia. Wikimedia Foundation, Inc. [Ηλεκτρονικό]. [Πρόσβαση 07 Jun 2012].
- [9] Wikipedia contributors, «Pattern recognition,» Wikipedia, the free encyclopedia. Wikimedia Foundation, Inc. [Ηλεκτρονικό]. [Πρόσβαση 29 May 2012].
- [10] A. P. Heath και L. E. Kaviraki, «Computational challenges in systems biology,» *Computer Science Review*, τόμ. 3, 2009.
- [11] J. M. Bower και H. Bolouri, Computational modeling of genetic and biochemical networks, MIT press, 2001.
- [12] A. Finkelstein, J. Hetherington, O. Margoninski, P. Saffrey, R. Seymour και A. Warner, «Computational challenges of systems biology,» *Computer*, τόμ. 37, 2004.
- [13] M. Tomita, «Whole-cell simulation: a grand challenge of the 21st century,» *TRENDS in Biotechnology*, τόμ. 19, 2001.

- [14] B. D. Davis και E. S. Mingioli, «Mutants of Escherichia coli requiring methionine or vitamin B12,» *Journal of Bacteriology*, τόμ. 60, 1950.
- [15] C. Dittamo και D. Cangelosi, «Optimized parallel implementation of gillespie's first reaction method on graphics processing units,» *Computer Modeling and Simulation, 2009. ICCMS'09. International Conference on*, 2009.
- [16] H. Li και L. Petzold, «Efficient parallelization of the stochastic simulation algorithm for chemically reacting systems on the graphics processing unit,» *International Journal of High Performance Computing Applications*, τόμ. 24, 2010.
- [17] D. T. Gillespie, «Stochastic simulation of chemical kinetics,» *Annu. Rev. Phys. Chem.*, τόμ. 58, 2007.
- [18] O. Hazapi και E. Manolakos, «Scalable FRM-SSA SoC design for the simulation of networks with thousands of biochemical reactions in real time,» *Field Programmable Logic and Applications (FPL), 2011 International Conference on*, 2011.
- [19] O. Hazapi, E. Logaras και E. S. Manolakos, «A soft IP core generating SoCs for the efficient stochastic simulation of large Biomolecular Networks using FPGAs,» *Electronics, Circuits and Systems (ICECS), 2012 19th IEEE International Conference on*, 2012.
- [20] M. Stepanova, F. Lin και V. Lin, «Design and development of a FPGA-based cascade Markov model for recognition of steroid hormone Response Elements,» *Parallel and Distributed Processing, 2008. IPDPS 2008. IEEE International Symposium on*, 2008.
- [21] H. Yamada, Y. Osana, T. Ishimori, T. Ooya, M. Yoshimi, Y. Nishikawa, A. Funahashi, N. Horoi, H. Amano και Y. Shibata, «A Modular Approach to Heterogeneous Biochemical Model Simulation on an FPGA,» *Reconfigurable Computing and FPGAs, 2009. ReConFig'09. International Conference on*, 2009.
- [22] M. Yoshimi, Y. Osana, T. Fukushima και H. Amano, «Stochastic simulation for biochemical reactions on FPGA,» *Field Programmable Logic and Application*, 2004.
- [23] M. Yoshimi, Y. Osana, Y. Iwaoka, Y. Nishikawa, T. Kojima, N. Hiroi, Y. Shibata, N. Iwanaga και H. Kitano, «An FPGA implementation of high throughput stochastic simulator for large-scale biochemical systems,» *Field Programmable Logic and Applications, 2006. FPL'06. International Conference on*, 2006.
- [24] N. Azizi, I. Kuon, A. Egier, A. Darabiha και P. Chow, «Reconfigurable molecular dynamics simulator,» *Field-Programmable Custom Computing Machines, 2004. FCCM 2004. 12th Annual IEEE Symposium on*, 2004.
- [25] J. F. Keane, C. Bradley και C. Ebeling, «A compiled accelerator for biological cell signaling simulations,» *Proceedings of the 2004 ACM/SIGDA 12th international symposium on Field programmable gate arrays*,

2004.

- [26] L. Macchiarulo, «A massively parallel implementation of gillespie algorithm on fpgas,» *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, 2008.
- [27] Y. Osana, T. Fukushima και H. Amano, «Implementation of recsip: a reconfigurable cell simulation platform,» *Field Programmable Logic and Application*, 2003.
- [28] Schneider και M. Victoria, *In Silico Systems Biology*, Humana Press, 2013, pp. 182-185.
- [29] M. Hucka, A. Finney, H. M. Sauro, H. Bolouri, J. Doyle, H. Kitano, A. Arkin, B. Bornstein, D. Bray και A. Cornish-Bowden, «The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models,» *Bioinformatics*, τόμ. 19, 2003.
- [30] Wikipedia contributors, «XML,» Wikipedia, the free encyclopedia. Wikimedia Foundation, Inc. [Ηλεκτρονικό]. [Πρόσβαση 01 Jun 2014].
- [31] C. J. Myers, «Engineering genetic circuits,» CRC Press, 2011.
- [32] D. J. Wilkinson, *Stochastic Modelling for Systems Biology*, CRC press, 2011.
- [33] D. J. Wilkinson, «Stochastic modelling for systems biology,» CRC press, 2011.
- [34] C. J. Myers, «Engineering genetic circuits,» CRC Press, 2011.
- [35] D. J. Wilkinson, «Stochastic modelling for systems biology,» CRC press, 2011.
- [36] D. T. Gillespie, «Exact stochastic simulation of coupled chemical reactions,» *The journal of physical chemistry*, τόμ. 81, 1977.
- [37] M. A. Gibson και J. Bruck, «Efficient exact stochastic simulation of chemical systems with many species and many channels,» *The journal of physical chemistry A*, τόμ. 104, 2000.
- [38] Xilinx, «Xilinx,» Xilinx Inc, [Ηλεκτρονικό]. Available: <http://www.xilinx.com/>. [Πρόσβαση 01 Jun 2013].
- [39] F. Horn και R. Jackson, «General mass action kinetics,» *Archive for Rational Mechanics and Analysis*, τόμ. 47, 1972.
- [40] N. H. Beebe, «A Bibliography of IEEE Standards,» 2013.
- [41] Electronic Industries Association. Engineering Dep, Interface between data terminal equipment and

- data communication equipment employing serial binary data interchange, Electronic Industries Association, Engineering Dept., 1969.
- [42] Metcalfe, Robert M and Boggs και R. David, «Ethernet: distributed packet switching for local computer networks,» *Communications of the ACM*, τόμ. 19, pp. 395-404, 1976.
- [43] D. Anderson, T. Shanley και R. Budruk, PCI express system architecture, Addison-Wesley Professional επιμ., Addison-Wesley Professional, 2004.
- [44] Xilinx, «ISE Design Suite,» Xilinx Inc, [Ηλεκτρονικό]. Available: <http://www.xilinx.com/products/design-tools/ise-design-suite.html>. [Πρόσβαση 01 Jul 2011].
- [45] Xilinx, «Running the Generate Programming File Process for FPGAs,» Xilinx Inc, [Ηλεκτρονικό]. Available: http://www.xilinx.com/support/documentation/sw_manuals/xilinx11/ise_p_generate_fpga_programming_file.htm. [Πρόσβαση 01 Jun 2013].
- [46] Xilinx, «iMPact Overview,» Xilinx Inc, [Ηλεκτρονικό]. Available: http://www.xilinx.com/itp/xilinx10/isehelp/pim_c_overview.htm. [Πρόσβαση 01 Jul 2012].
- [47] Vermeulen, Bart and Waayers, Tom and Bakker και Sjaak, «IEEE 1149.1-compliant access architecture for multiple core debug on digital system chips,» *Test Conference, 2002. Proceedings. International*, pp. 55-63, 2002.
- [48] Wikipedia contributors, «Pseudorandom number generator,» Wikipedia, the free encyclopedia. Wikimedia Foundation, Inc. [Ηλεκτρονικό]. [Πρόσβαση 01 Jul 2011].
- [49] Wikipedia contributors, «Comma-separated values,» Wikipedia, the free encyclopedia. Wikimedia Foundation, Inc. [Ηλεκτρονικό]. [Πρόσβαση 01 Jul 2011].
- [50] T. Henriksson και P. Van der Wolf, «Ttl hardware interface: A high-level interface for streaming multiprocessor architectures,» *Proceedings of the 2006 IEEE/ACM/IFIP Workshop on Embedded Systems for Real Time Multimedia*, 2006.
- [51] A. AMBA, «Specification Rev 2.0,» {ARM Limited, 1999.
- [52] M. V. Schneider, In Silico Systems Biology, Humana Press, 2013.
- [53] S. Hoops, S. Sahle, R. Gauges, C. Lee, J. Pahle, S. Natalia, M. Singhal, L. Xu, P. Mendes και U. Kummer, «COPASI—a complex pathway simulator,» *Bioinformatics*, τόμ. 22, 2006.
- [54] Xilinx, «Xilinx CORE Generator System,» Xilinx Inc. [Ηλεκτρονικό]. [Πρόσβαση 01 Jul 2011].

- [55] M. Matsumoto και T. Nishimura, «Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator,» *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, τόμ. 8, αρ. ACM, 1998.
- [56] «Mersenne Twister, MT32,» HT-LAB, [Ηλεκτρονικό]. Available: <http://www.ht-lab.com/freecores/mt32/mersenne.html>. [Πρόσβαση 1 Jul 2014].
- [57] Y. Cao, D. T. Gillespie και L. R. Petzold, «Efficient step size selection for the tau-leaping simulation method,» *The Journal of chemical physics*, τόμ. 124, 2006.
- [58] Wikipedia contributors, «Verification and validation,» Wikipedia, the free encyclopedia. Wikimedia Foundation, Inc. [Ηλεκτρονικό]. [Πρόσβαση 01 Jul 2013].
- [59] Xilinx, «ISE Simulator (ISim),» Xilinx Inc, [Ηλεκτρονικό]. Available: <http://www.xilinx.com/tools/isim.htm>. [Πρόσβαση 01 Jun 2013].
- [60] Xilinx, «Behavioral Simulation,» Xilinx Inc, [Ηλεκτρονικό]. Available: http://www.xilinx.com/support/documentation/sw_manuels/xilinx11/platform_studio/ps_c_sim_behavioral_simulation.htm. [Πρόσβαση 01 Jun 2013].
- [61] Xilinx, «Timing Simulation,» Xilinx Inc, [Ηλεκτρονικό]. Available: http://www.xilinx.com/support/documentation/sw_manuels/xilinx11/ise_c_simulation_timing.htm. [Πρόσβαση 01 Jun 2013].
- [62] Y. Zhou, J. Liepe, X. Sheng, M. Stumpf και C. Barnes, «GPU accelerated biochemical network simulation,» *Bioinformatics*, τόμ. 27, 2011.
- [63] D. J. Wilkinson, «Stochastic modelling for systems biology,» CRC press, 2011.
- [64] E. Logaras, O. Hazapi και E. Manolakos, «Python to accelerate embedded SoC design: a case study for systems biology,» *ACM Transactions on Embedded Computing Systems (TECS)*, τόμ. 13, 2014.
- [65] G. Klingbeil, R. Erban, M. Giles και P. K. Maini, «Stochsimgpu: parallel stochastic simulation for the systems biology toolbox 2 for matlab,» *Bioinformatics*, τόμ. 27, 2011.
- [66] Ullah, Mukhtar and Wolkenhauer και Olaf, «Stochastic approaches in systems biology,» *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, τόμ. 2, pp. 385-397, 2010.