



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ  
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ  
ΤΟΜΕΑΣ ΕΠΙΚΟΙΝΩΝΙΩΝ, ΗΛΕΚΤΡΟΝΙΚΗΣ & ΣΥΣΤΗΜΑΤΩΝ  
ΠΛΗΡΟΦΟΡΙΚΗΣ

## Αρχιτεκτονικές και Μηχανισμοί διασφάλισης Ποιότητας Υπηρεσιών σε Υπηρεσιοστρεφείς Υποδομές

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Ανδρέας Ι. Μενύχτας

Αθήνα, Μάιος 2009





ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ  
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ  
ΤΟΜΕΑΣ ΕΠΙΚΟΙΝΩΝΙΩΝ, ΗΛΕΚΤΡΟΝΙΚΗΣ & ΣΥΣΤΗΜΑΤΩΝ  
ΠΛΗΡΟΦΟΡΙΚΗΣ

## Αρχιτεκτονικές και Μηχανισμοί διασφάλισης Ποιότητας Υπηρεσιών σε Υπηρεσιοστρεφείς Υποδομές

### ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Ανδρέας Ι. Μενύχτας

**Συμβουλευτική Επιτροπή :** Θεοδώρα Βαρβαρίγου  
Εμμανουήλ Πρωτονοτάριος  
Ελευθέριος Καγιάφας

Εγκρίθηκε από την επταμελή εξεταστική επιτροπή την 14<sup>η</sup> Μαΐου 2009.

.....  
Θεοδώρα Βαρβαρίγου  
Καθηγήτρια Ε.Μ.Π.

.....  
Εμμανουήλ Πρωτονοτάριος  
Καθηγητής Ε.Μ.Π.

.....  
Ελευθέριος Καγιάφας  
Καθηγητής Ε.Μ.Π.

.....  
Γεώργιος Στασινόπουλος  
Καθηγητής Ε.Μ.Π.

.....  
Νεκτάριος Κοζύρης  
Καθηγητής Ε.Μ.Π.

.....  
Συμεών Παπαβασιλείου  
Καθηγητής Ε.Μ.Π.

.....  
Αναστάσιος Δουλάμης  
Καθηγητής Πολυτεχνείου  
Κρήτης

Αθήνα, Μάιος 2009

.....

Ανδρέας Ι. Μενύχτας

Διδάκτωρ Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © Ανδρέας Ι. Μενύχτας, 2009.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

*“Es ist nicht genug zu wissen, man muss auch anwenden; es ist nicht genug zu wollen, man muss auch tun.”*

*“Η γνώση δεν είναι αρκετή, πρέπει να εφαρμόζουμε. Η θέληση δεν είναι αρκετή, πρέπει να πράττουμε.”*

--Goethe, Johann Wolfgang Von--

*Η σελίδα αυτή είναι σκόπιμα λευκή*

## **Πρόλογος**

*Η διδακτορική διατριβή που παρουσιάζεται στις επόμενες σελίδες εκπονήθηκε από τον Οκτώβριο του 2005 μέχρι το Μάιο του 2009, στο εργαστήριο Τηλεπικοινωνιών του τομέα Επικοινωνιών, Ηλεκτρονικής και Συστημάτων Πληροφορικής, στη Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του Εθνικού Μετσόβιου Πολυτεχνείου.*

*Κατά τη διάρκεια της εκπόνησης της διατριβής, είχα την ευκαιρία να ασχοληθώ με αρκετά ενδιαφέροντα επιστημονικά θέματα που αφορούν κυρίως στους τομείς του αρχιτεκτονικού σχεδιασμού και διαχείρισης υπηρεσιοστρεφών υποδομών και διασφάλισης ποιότητας υπηρεσιών και να αποκτήσω πολύτιμη εμπειρία και γνώσεις.*

*Θα ήθελα να ευχαριστήσω από τα βάθη της καρδιάς μου, την καθηγήτρια μου κ. Θεοδώρα Βαρβαρίγου για το ενδιαφέρον που έδειξε, για τις πολύτιμες συμβουλές της και για την ιδιαίτερη στήριξη που μου παρείχε κατά την διάρκεια αυτής της πορείας μου, καθώς επίσης τους καθηγητές της τριμελούς συμβουλευτικής επιτροπής κ.κ. Εμμανουήλ Πρωτονοτάριο και Ελευθέριο Καγιάφα.*

*Επίσης, θα ήθελα να ευχαριστήσω όλους τους συναδέλφους μου στην ερευνητική ομάδα με τους οποίους συνεργάστηκα άμογα και επιτυχώς όλα αυτά τα χρόνια. Ιδιαίτερες ευχαριστίες ωστόσο θα ήθελα να απευθύνω στους στενούς μου συνεργάτες Δημοσθένη Κυριαζή, Κωνσταντίνο Τσερπέ, Κλεοπάτρα Κωνσταντέλη, Δημήτρη Χάλκο, Κωστή Χριστογιάννη, Γιώργο Κουσιουρή, Αναστάσιο Δουλάμη και Νικόλαο Δουλάμη με τους οποίους μοιραστήκαμε τις πάρα πολλές ώρες της ερευνητικής εργασίας.*

*Τέλος, ένα μεγάλο ευχαριστώ οφείλω στους δικούς μου ανθρώπους που με στήριξαν σε αυτήν την πορεία και ιδιαίτερα στους γονείς μου που πίστεψαν σε εμένα και τις επιλογές μου.*

*Ανδρέας Ι. Μενύχτας*

*Μάιος 2009*

*Η σελίδα αυτή είναι σκόπιμα λευκή*



## *Πίνακας Περιεχομένων*

Περίληψη.....	1
Abstract .....	3
1 Εισαγωγή.....	5
1.1 Το ερευνητικό πλαίσιο και οι τεχνολογικές προκλήσεις της διατριβής.....	5
1.2 Το όραμα και οι καινοτομίες της διατριβής.....	9
1.3 Πλέγματα, υπηρεσιοστρεφείς υποδομές και ποιότητα υπηρεσιών.....	10
1.3.1 Ορισμός του πλέγματος.....	10
1.3.2 Υπηρεσιοστρεφείς υποδομές.....	12
1.3.3 Ποιότητα υπηρεσιών .....	13
1.4 Οργάνωση του εγγράφου.....	14
2 Οι υπηρεσιοστρεφείς υποδομές και το περιβάλλον πλέγματος .....	17
2.1 Ιστορικά στοιχεία - Από το διαδίκτυο στο πλέγμα.....	17
2.2 Χαρακτηριστικά.....	21
2.3 Οι γενιές των συστημάτων πλέγματος.....	23
2.4 Ορολογία σχετική με πλέγματα .....	25
2.4.1 Το meta-computing .....	25
2.4.2 Μη υπολογιστικοί πόροι.....	26
2.4.3 Εικονοποίηση .....	27
2.4.4 Υπηρεσίες δικτύου .....	28
2.4.5 Εικονικοί οργανισμοί .....	29
2.5 Ταξινόμηση των πλεγμάτων.....	31
2.6 Αρχιτεκτονική σχεδίαση του πλέγματος .....	33
2.6.1 Απαιτήσεις ασφάλειας.....	34
2.6.2 Ευαισθησία δεδομένων .....	35
2.6.3 Μέγιστες απαιτήσεις σε επεξεργαστική ισχύ .....	36
2.6.4 Αποθήκευση δεδομένων.....	36
2.6.5 Διαθέσιμο εύρος ζώνης στο διαδίκτυο .....	37
2.6.6 Υπάρχοντες πόροι .....	38
2.6.7 Πόροι ειδικού σκοπού .....	40
2.6.8 Δυνατότητες συνεργασίας.....	41
2.7 Σύγκριση των τεχνολογιών πλέγματος με άλλες τεχνολογίες.....	41
2.8 Τα πλεονεκτήματα του πλέγματος.....	43
2.9 Τομείς που ευνοούνται από την τεχνολογία πλέγματος.....	44
3 Υπηρεσιοστρεφείς αρχιτεκτονικές και πλέγματα: Η σημερινή κατάσταση.....	47
3.1 Υπηρεσιοστρεφής αρχιτεκτονική .....	47
3.1.1 Γενικά για τις υπηρεσίες.....	48

3.1.2	Διαδικτυακές υπηρεσίες .....	49
3.1.3	Web Services Description Language (WSDL).....	51
3.1.4	Simple Object Access Protocol (SOAP) .....	53
3.2	Αρχιτεκτονική ανοιχτών υπηρεσιών πλέγματος.....	56
3.2.1	Στόχοι της αρχιτεκτονικής.....	56
3.2.2	Περιγραφή της αρχιτεκτονικής .....	57
4	Ποιότητα υπηρεσιών και επιχειρηματικά μοντέλα .....	61
4.1	Ποιότητα υπηρεσιών: Η σημερινή κατάσταση.....	61
4.1.1	Περιγραφή αίτησης για ποιότητα υπηρεσιών.....	61
4.1.2	Περιγραφή υπηρεσιών.....	62
4.1.3	Προσαρμοστική ποιότητα υπηρεσιών .....	62
4.1.4	Δέσμευση πόρων και δρομολόγηση .....	63
4.1.5	Αξιολόγηση παροχής ποιότητας υπηρεσιών .....	63
4.2	Επιχειρηματικά μοντέλα .....	63
4.2.1	Τι είναι τα επιχειρηματικά μοντέλα; .....	64
4.2.2	Οι ρόλοι σε ένα επιχειρηματικό μοντέλο .....	66
4.2.3	Επιχειρηματικές απαιτήσεις .....	69
4.3	Αρχές σχεδιασμού της αρχιτεκτονικής.....	73
4.3.1	Κύριες αρχές σχεδιασμού.....	75
4.3.2	Δευτερεύουσες αρχές σχεδιασμού .....	83
4.4	Εφαρμογή σε ένα καινοτόμο σενάριο χρήσης: Υπηρεσίες πολυμέσων σε περιβάλλον πλέγματος.....	85
4.4.1	Ποιότητα υπηρεσιών και επιχειρηματικά μοντέλα για εφαρμογές πολυμέσων σε περιβάλλον πλέγματος.....	89
4.5	Καινοτόμος υλοποίηση υπηρεσιοστρεφών υποδομών για τη διασφάλιση ποιότητας υπηρεσιών.....	92
5	Περιβάλλοντα πλέγματος επόμενης γενιάς .....	93
5.1	Εφαρμογή πολυμέσων σε περιβάλλον πλέγματος.....	93
5.1.1	Χαρακτηριστικά γνωρίσματα πλέγματος επόμενης γενιάς.....	94
5.1.2	Συμφωνίες επιπέδου υπηρεσιών.....	95
5.1.3	Συμφωνίες επιπέδου υπηρεσιών για εφαρμογές πολυμέσων .....	96
5.1.4	Ο ρόλος των ροών εργασιών .....	102
5.2	Γενικό διάγραμμα ενσωμάτωσης.....	102
5.3	Αρχιτεκτονική πλέγματος για εφαρμογές πολυμέσων .....	103
5.3.1	Ενσωμάτωση στοιχείων πλέγματος.....	103
5.3.2	Ακολουθία λειτουργίας .....	109
5.4	Πειράματα αξιολόγησης .....	115

5.4.1	Πείραμα 1: Αξιολόγηση επιπέδου ποιότητας υπηρεσιών στην ψηφιακή εφαρμογή μέσω...	115
5.4.2	Αποτελέσματα	115
5.4.3	Πείραμα 2: Δρομολόγηση ροών εργασιών με βάση τις συμφωνίες επιπέδου υπηρεσιών...	124
6	Λειτουργία του πλέγματος σε πραγματικό χρόνο	127
6.1	Παροχή ποιότητας υπηρεσιών σε υποδομές πραγματικού χρόνου	130
6.2	Αρχιτεκτονική προσέγγιση	133
6.2.1	Αποθήκη συμφωνιών επιπέδου υπηρεσιών	135
6.2.2	Αποθήκη πολιτικών διαχείρισης	135
6.2.3	Έλεγχος πραγματικού χρόνου	136
6.2.4	Τμήματα σύστασης	137
6.2.5	Επόπτης ποιότητας υπηρεσιών	137
6.3	Αλληλεπιδράσεις βασικών δομικών στοιχείων	138
6.4	Το σενάριο της τρισδιάστατης απεικόνισης	141
6.5	Συμπεράσματα	147
7	Υπηρεσιοστρεφείς υποδομές για διαδραστικές εφαρμογές πραγματικού χρόνου	149
7.1	Αρχές σχεδιασμού της αρχιτεκτονικής	149
7.1.1	Υπηρεσιοστρεφείς υποδομή	153
7.1.2	Απόκριση σε πραγματικό χρόνο	154
7.2	Συμφωνίες επιπέδου υπηρεσιών	156
7.2.1	Διαχείριση συμφωνιών επιπέδου υπηρεσιών	156
7.2.2	Τύποι συμφωνιών επιπέδου υπηρεσιών και η σχέση τους με την αλυσίδα προστιθέμενης αξίας	157
7.2.3	Καινοτομίες στη διαχείριση συμφωνιών επιπέδου υπηρεσιών	159
7.2.4	Εικονοποίηση των πόρων και διασφάλιση της ποιότητας υπηρεσιών	164
7.3	Καινοτόμος ορισμός της αρχιτεκτονικής	165
7.4	Καινοτόμες υπηρεσίες πλαισίου	168
7.4.1	Επισκόπηση υπηρεσιών πλαισίου	169
7.4.2	Φάσεις λειτουργίας	184
7.4.3	Διεπαφές υπηρεσιών πλαισίου	186
7.5	Συμπεράσματα	188
8	Καινοτόμος ενσωμάτωση εφαρμογών βιοϊατρικής σε περιβάλλον πλέγματος	189
8.1	Ποιότητα υπηρεσίας σε περιβάλλον πλέγματος	190
8.1.1	Βασικά στοιχεία αρχιτεκτονικής για τη διασφάλιση ποιότητας υπηρεσιών	192
8.2	Εκτίμηση υπολογιστικού φόρτου	194
8.3	Εκτίμηση απόδοσης πόρων	195
8.3.1	Καταγραφή επιπέδου ποιότητας υπηρεσιών και αξιοποίηση της	197

8.4	Δρομολόγηση εργασιών .....	198
8.4.1	Διαχείριση πόρων και δρομολόγηση.....	200
8.4.2	Δρομολόγηση στο μεσολογισμικό .....	201
8.5	Λεπτομέρειες υλοποίησης .....	205
8.6	Καινοτόμος αρχιτεκτονική προσέγγιση.....	206
8.6.1	Παρουσίαση.....	209
8.6.2	Λειτουργικότητα.....	209
8.6.3	Βάση Δεδομένων.....	210
8.6.4	Πλέγμα.....	210
8.7	Αποτελέσματα .....	210
8.7.1	Σενάριο χρήσης: Ακτινοθεραπεία .....	210
8.7.2	Σενάριο χρήσης: Χημειοθεραπεία.....	213
8.8	Συμπεράσματα και μελλοντικά σχέδια.....	214
9	Συμπεράσματα και μελλοντικές εργασίες .....	215
9.1	Συμπεράσματα .....	215
9.2	Μελλοντικές εργασίες .....	217
10	Συντομογραφίες.....	219
11	Βιβλιογραφικές Αναφορές .....	221

## Σχήματα

Σχήμα 1: Αναδυόμενες τεχνολογίες (2005) .....	7
Σχήμα 2: Αναδυόμενες τεχνολογίες (2008) .....	7
Σχήμα 3: Αρχιτεκτονική Globus .....	21
Σχήμα 4: Δομή υπηρεσιοστρεφούς αρχιτεκτονικής.....	49
Σχήμα 5: Ανταλλαγή ρόλων μεταξύ Web services κατά τη διάρκεια μιας επικοινωνίας.....	51
Σχήμα 6: Χρήση SOAP Πρωτοκόλλου σε εφαρμογές.....	53
Σχήμα 7: Βασική δομή μηνύματος SOAP.....	55
Σχήμα 8: Αρχιτεκτονική ανοιχτών υπηρεσιών πλέγματος.....	57
Σχήμα 9: Η αντιστοιχία των απαιτήσεων των καταναλωτών με τους πόρους των υπηρεσιών μέσω της χρήσης συμφωνιών επιπέδου υπηρεσιών .....	77
Σχήμα 10: Κοινή εκμετάλλευση πόρων .....	81
Σχήμα 11: Σύνθεση πόρων .....	81
Σχήμα 12: Ενθυλάκωση πόρων .....	81
Σχήμα 13: Ο σχεδιασμός και το αποτέλεσμα τρισδιάστατων απεικονίσεων .....	87
Σχήμα 14: Η διαδικασία δημιουργίας τρισδιάστατων απεικονίσεων.....	87
Σχήμα 15: Διάγραμμα ενσωμάτωσης.....	103
Σχήμα 16: Διάγραμμα ενσωμάτωσης για εφαρμογές πολυμέσων .....	105
Σχήμα 17: Παράδειγμα μηνύματος ενημέρωσης .....	109
Σχήμα 18: Δημιουργία προτύπων ροών εργασιών, συμφωνιών επιπέδου υπηρεσιών και πολιτικών διαχείρισης .....	109
Σχήμα 19: Λειτουργίες καταναλωτή .....	111
Σχήμα 20: Αρχιτεκτονική και αλληλουχία χρήσης .....	113
Σχήμα 21: Αξιολόγηση για χρήση 40% .....	117
Σχήμα 22: Αξιολόγηση για χρήση 80% .....	119
Σχήμα 23: Απόδοση για διαφορετικό αριθμό καταγεγραμμένων δεδομένων.....	121
Σχήμα 24: Απόδοση για διαφορετική χρήση συμφωνιών επιπέδου υπηρεσιών .....	123
Σχήμα 25: Διάγραμμα Αρχιτεκτονικής .....	133
Σχήμα 26: Διάγραμμα Αλληλεπίδρασης .....	139
Σχήμα 27: Συμφωνίες επιπέδου υπηρεσιών και πολιτικές διαχείρισης για την εφαρμογή 3D Rendering .....	141
Σχήμα 28: Επιδόσεις Εφαρμογής Πολιτικών .....	145
Σχήμα 29: Συνολική Επίδοση Εφαρμογής Πολιτικών .....	145
Σχήμα 30: Υπηρεσιοστρεφής υποδομή πραγματικού χρόνου.....	151
Σχήμα 31: Οι τύποι συμφωνιών επιπέδου υπηρεσιών .....	159
Σχήμα 32: Τα πλεονεκτήματα της υποδομής .....	161
Σχήμα 33: Η γενική αρχιτεκτονική .....	165
Σχήμα 34: Η αρχιτεκτονική των υπηρεσιών πλαισίου .....	171

Σχήμα 35: Υπόμνημα αρχιτεκτονικής υπηρεσιών πλαισίου .....	171
Σχήμα 36: Εργαλεία υποστήριξης εφαρμογής .....	173
Σχήμα 37: Διαφήμιση και αναζήτηση .....	177
Σχήμα 38: Διαχείριση συμφωνιών επιπέδου υπηρεσιών.....	179
Σχήμα 39: Διαχείριση ροών εργασιών .....	181
Σχήμα 40: Αποτελέσματα εκτίμησης απόδοσης πόρων.....	195
Σχήμα 41: Αναπαράσταση της υλοποίησης και λειτουργίας της εφαρμογής .....	207

## ***Πίνακες***

Πίνακας 1: Παράδειγμα Shader SLA .....	97
Πίνακας 2: Παράδειγμα Texture SLA.....	97
Πίνακας 3: Προδιαγραφές πρότυπης μηχανής .....	99
Πίνακας 4: Κατηγορίες πολυπλοκότητας.....	99
Πίνακας 5: Παράδειγμα Rendering SLA .....	101
Πίνακας 6: Στατιστικά χρόνων εκτέλεσης στο πλέγμα.....	212





## **Περίληψη**

Οι υπηρεσιοστρεφείς αρχιτεκτονικές (Service Oriented Architectures – SOA) αποτελούν αδιαμφισβήτητα την επικρατούσα τάση για την ανάπτυξη των σύγχρονων συστημάτων πληροφορικής και επικοινωνιών. Επιπρόσθετα, χαρακτηριστικό γνώρισμα όλων των σύγχρονων συστημάτων, και ειδικότερα αυτών με επιχειρηματικό προσανατολισμό, είναι η ανάγκη για διασφάλιση του επιπέδου παροχής ποιότητας υπηρεσιών (Quality of Service – QoS) για τους χρήστες τους, κάτι που καθορίζει τόσο την αποδοχή τους στο ευρύ κοινό όσο και την οικονομική βιωσιμότητά τους.

Η διατριβή επικεντρώνεται στη μελέτη ολοκληρωμένων αρχιτεκτονικών και μηχανισμών με σκοπό τη διασφάλιση ποιότητας υπηρεσιών σε υπηρεσιοστρεφείς υποδομές (SOIs). Ο στόχος της διατριβής δεν είναι μόνο η διασφάλιση της ποιότητας υπηρεσιών η οποία μπορεί να ορίζεται γενικά ή μέσα από διεθνή πρότυπα αλλά και μεμονωμένα/ειδικά για κάθε υπηρεσία, όπως την αντιλαμβάνονται οι εμπλεκόμενες οντότητες κατά τη διαδικασία παροχής/χρήσης της και την εκφράζουν στη μεταξύ τους συμφωνία.

Καθοριστικό στοιχείο για το επίπεδο της ποιότητας υπηρεσιών και τον τρόπο που αυτή παρέχεται είναι το επιχειρηματικό μοντέλο στο οποίο εντάσσεται η υποδομή και οι υπηρεσίες που προσφέρονται. Στο πλαίσιο αυτό αναλύεται, σχεδιάζεται και υλοποιείται μια σειρά από ολοκληρωμένες υπηρεσιοστρεφείς αρχιτεκτονικές λύσεις και μηχανισμούς που αντιμετωπίζουν τις προκλήσεις για την διασφάλιση της ποιότητας υπηρεσιών στον ευρύ τομέα των υπηρεσιοστρεφών υποδομών και των πλεγμάτων επόμενης γενιάς. Ιδιαίτερη έμφαση δίνεται σε εφαρμογές βιοϊατρικής, όπου προτείνονται και αξιολογούνται μηχανισμοί που διασφαλίζουν την ποιότητα υπηρεσιών μέσα από δυναμική δρομολόγηση και αποδοτικό υπολογισμό φόρτου εργασίας και χωρητικότητας. Παράλληλα μελετήθηκαν οι επιδόσεις των υποδομών και των μηχανισμών, τόσο από την πλευρά του

χρήστη όσο και από την πλευρά του παρόχου υπηρεσιών, αναφορικά με επιχειρηματικές εφαρμογές όπως οι εφαρμογές πολυμέσων.

Τέλος μελετήθηκε και παρουσιάζεται μια καινοτόμος αρχιτεκτονική προσέγγιση που αναπτύχθηκε με σκοπό να καλύψει τις ανάγκες διαδραστικών εφαρμογών πραγματικού χρόνου, οι οποίες εκτελούνται σε εικονικά περιβάλλοντα υπηρεσιοστρεφών υποδομών.

## ***Abstract***

Service Oriented Architectures (SOA) constitute a prevailed tendency for the development of modern information and communication systems. Additionally, an important characteristic of modern systems, specifically for these with business orientation, is the need for guaranteeing the Quality of Service (QoS) levels for the end users, which is the key factor for their adoption from the community and also for their commercial viability.

The thesis focuses on the study of architectures and mechanisms that support the guaranteeing of Quality of Service in Service Oriented Infrastructures (SOI). The objective of the thesis is not only the guaranteeing of Quality of Service that can be defined in a generic way or based on international standards, but also for QoS that is individually/specifically defined for each service and realized by the entities involved in the process of using/providing the service and is agreed in their contract.

The business model, through which the infrastructure offers the services, has a considerable effect on the level of Quality of Service and the way that is provided. In this frame, several object oriented architectural solutions and mechanisms, which tackle the challenges of the guaranteeing of Quality of Services in the wide sector of Object Oriented Infrastructures and Next Generation Grids, are analyzed, designed and developed.

Furthermore we focused on biomedical applications and to this direction, mechanisms that guarantee the Quality of Services through dynamic scheduling, workload and capacity estimation, are proposed and evaluated. The architectures and the mechanisms on both user and service provider side have been also analysed for business applications such as the digital media applications.

Finally, an innovative architectural approach that covers the requirements of interactive real-time applications that are deployed and executed in virtual environments of Object Oriented Infrastructures is presented and evaluated.

# 1

## *Εισαγωγή*

Στο εισαγωγικό αυτό κεφάλαιο παρουσιάζουμε το ερευνητικό πλαίσιο της διατριβής, τις τεχνολογικές προκλήσεις και τα προβλήματα που αντιμετωπίζει καθώς και το όραμα και τις καινοτομίες που προέκυψαν. Επιπρόσθετα παραθέτουμε τις βασικές τεχνολογίες που σχετίζονται με τη διατριβή και τέλος τον τρόπο με τον οποίο οργανώθηκε.

### *1.1 Το ερευνητικό πλαίσιο και οι τεχνολογικές προκλήσεις της διατριβής*

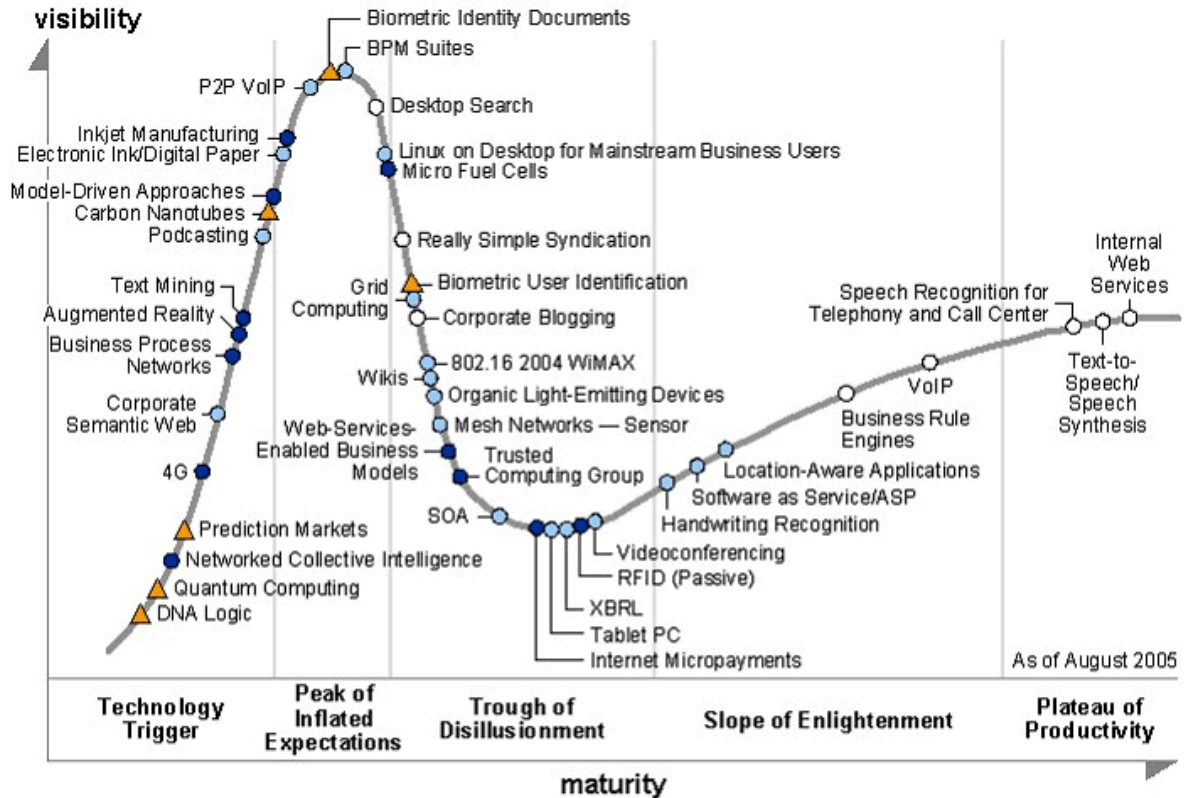
Οι τεχνολογίες πληροφορικής και επικοινωνιών διαδραματίζουν σήμερα ένα βασικό ρόλο στη ζωή του σύγχρονου ανθρώπου. Η επιρροή, άλλοτε άμεση και άλλοτε έμμεση, που έχουν αυτές οι τεχνολογίες είναι καθοριστική στο πως αναπτύσσεται ο σύγχρονος τρόπος ζωής, κυρίως γιατί γύρω από αυτές τις τεχνολογίες έχουν αναπτυχθεί επιχειρηματικά μοντέλα όπως η κοινωνική δικτύωση (social networks) και τα ηλεκτρονικά καταστήματα, τα οποία είναι κομμάτι πλέον της καθημερινότητάς μας. Αλλά συμβαίνει και το αντίθετο, καινοτόμα επιχειρηματικά μοντέλα να δημιουργούν νέες τεχνολογικές τάσεις προκειμένου να τα καταστήσουν εφικτά, είτε λύνοντας προβλήματα που οι προηγούμενες τεχνολογίες δεν μπορούσαν να αντιμετωπίσουν, είτε δημιουργώντας νέες ανάγκες.

Αντιλαμβανόμαστε λοιπόν πως ο τομέας των τεχνολογιών πληροφορικής και επικοινωνιών είναι ευρύς και το κυριότερο, ευμετάβλητος. Τεχνολογίες που δε συνοδεύονται από βιώσιμα επιχειρηματικά μοντέλα καθίστανται σύντομα ξεπερασμένες,

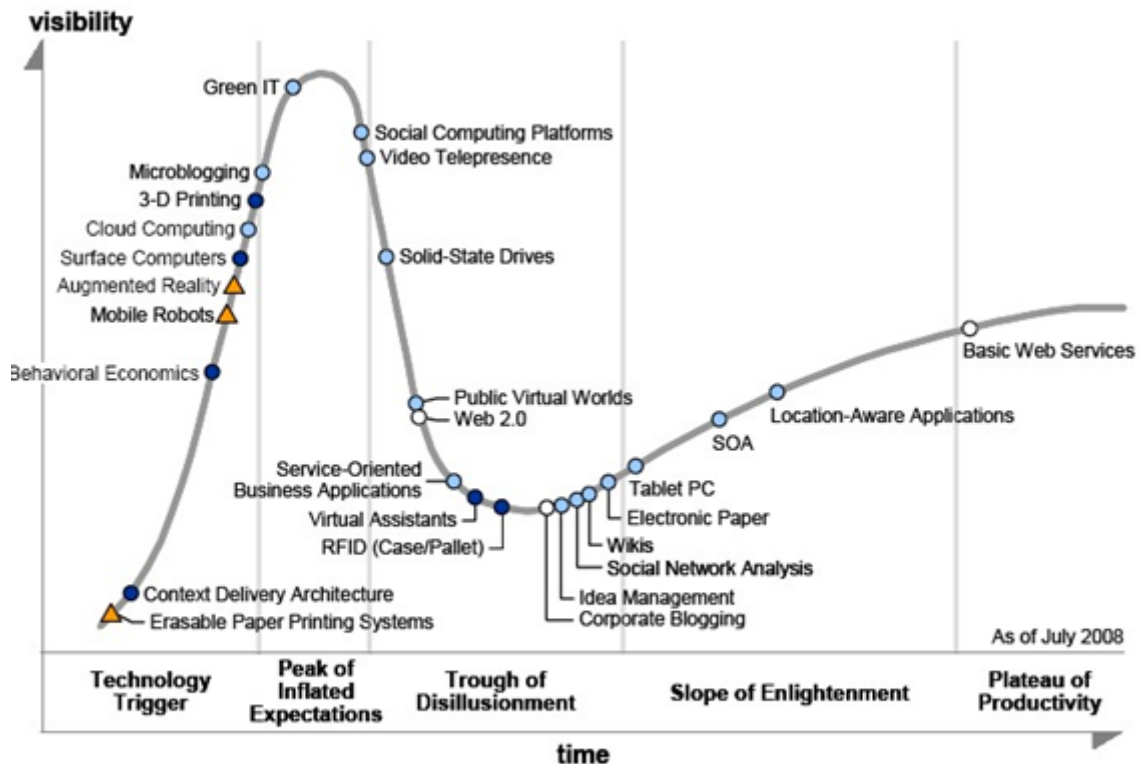
ενώ αυτές που έχουν ενδιαφέρον επιχειρηματικό υπόβαθρο, διαδίδονται και εξελίσσονται πολύ γρήγορα. Η διατριβή εντάσσεται στον ιδιαίτερο ενδιαφέρον τομέα των κατανεμημένων υπολογιστικών συστημάτων (Distributed Computing) και συγκεκριμένα των τεχνολογιών πλέγματος (Grid) [37], στον οποίο παρατηρούμε όλα τα παραπάνω χαρακτηριστικά.

Οι τεχνολογίες πλέγματος δεν είναι κάτι καινούριο, αλλά εξελίσσονται σε διάφορες παραλλαγές αρκετά χρόνια. Και σε αυτό τον τομέα τα επιχειρηματικά μοντέλα δημιούργησαν νέες τάσεις, με κυριότερη την υλοποίηση των υποδομών τους ακολουθώντας υπηρεσιοστρεφείς αρχιτεκτονικές (Service Oriented Architecture – SOA) [65].

Στα ακόλουθα διαγράμματα τύπου Hype Cycle [40] από την Gartner Research για τις αναδυόμενες τεχνολογίες παρατηρούμε ακριβώς την εξέλιξη των τεχνολογιών πλέγματος και υπηρεσιοστρεφών αρχιτεκτονικών. Το πρώτο διάγραμμα αφορά το έτος 2005 ενώ το δεύτερο το έτος 2008.



Σχήμα 1: Αναδύομενες τεχνολογίες (2005)



Σχήμα 2: Αναδύομενες τεχνολογίες (2008)

*Η σελίδα αυτή είναι σκόπιμα λευκή*



Οι τεχνολογίες πλέγματος στο έτος 2005 είναι ιδιαίτερα δημοφιλείς όπως και οι τεχνολογίες υπηρεσιοστρεφών αρχιτεκτονικών. Στο έτος 2008 βλέπουμε πως οι τεχνολογίες πλέγματος δεν εμφανίζονται (τουλάχιστον άμεσα) και «κρύβονται» πίσω από τις υπηρεσιοστρεφείς υποδομές, οι οποίες καθίστανται ακόμη πιο ώριμες, αλλά και από μια νέα τάση, το Cloud Computing [78].

Κοινό χαρακτηριστικό σε όλα τα παραπάνω είναι η **ποιότητα υπηρεσιών** και ο τρόπος με τον οποίο αυτή διασφαλίζεται. Στα πλαίσια της διατριβής παρακολουθήσαμε στενά τις τεχνολογικές εξελίξεις αναφορικά με την ποιότητα υπηρεσιών διασφαλίζοντας ότι τα αποτελέσματά που παρουσιάζονται είναι εφαρμόσιμα στις νέες υποδομές και τεχνολογίες.

## ***1.2 Το όραμα και οι καινοτομίες της διατριβής***

Η ανάγκη για ποιότητα υπηρεσίας είναι αναπόσπαστο στοιχείο των σύγχρονων επιχειρηματικών μοντέλων. Άμεσα ή έμμεσα, η παροχή ποιότητα υπηρεσιών παίζει καθοριστικό ρόλο σε ένα επιχειρηματικό μοντέλο, κάτι που είναι ακόμη πιο έντονο όταν τα επιχειρηματικά μοντέλα εμπλέκουν αρκετές οντότητες, όπου η λειτουργία και το οικονομικό κέρδος της κάθε μιας εξαρτάται από την άλλη. Αντιλαμβανόμαστε λοιπόν πως δεν μπορεί να υπάρξει οικονομικά βιώσιμο καταναλωμένο σύστημα στο οποίο δεν διασφαλίζεται το επίπεδο παροχής ποιότητας υπηρεσιών μεταξύ των εμπλεκόμενων οντοτήτων (π.χ. μεταξύ του παρόχου και του καταναλωτή μιας υπηρεσίας).

Η διατριβή επικεντρώνεται σε αυτό ακριβώς τα ζήτημα, τη διασφάλιση της ποιότητας υπηρεσιών για τους χρήστες υπηρεσιοστρεφών υποδομών, κάτι που βρίσκει εφαρμογή από τα πλέγματα υπολογιστών μέχρι το νέο τεχνολογικό τομέα του Cloud Computing. Στο πλαίσιο της διατριβής ερευνήθηκαν και προτείνονται νέες καινοτόμες αρχιτεκτονικές λύσεις και μηχανισμοί για υπηρεσιοστρεφείς υποδομές με σκοπό:

- τη **διασφάλιση της ποιότητας υπηρεσιών** από αρχή έως τέλους στην υποδομή (end to end),
- την **περιγραφή της ποιότητας υπηρεσιών** με τρόπο που την αντιλαμβάνεται ο τελικός χρήστης, χωρίς να απαιτείται ιδιαίτερη γνώση για την υποδομή,
- τη δημιουργία **επιχειρηματικά βιώσιμων** υποδομών και
- την **επαναχρησιμοποίηση** των προτεινόμενων μηχανισμών και την **επεκτασιμότητα** τους σύμφωνα με τις ανάγκες του επιχειρηματικού μοντέλου.

### ***1.3 Πλέγματα, υπηρεσιοστρεφείς υποδομές και ποιότητα υπηρεσιών***

#### ***1.3.1 Ορισμός του πλέγματος***

Η τεχνολογία των υπολογιστών και των τηλεπικοινωνιών γνωρίζει την κορύφωση της τα τελευταία χρόνια καθώς αναπτύσσονται συνεχώς εφαρμογές που χαρακτηρίζονται από εξαιρετικές απαιτήσεις σε υπολογιστική ισχύ και χρόνο εκτέλεσης. Παράλληλα, εντείνεται η ζήτηση για πρόσβαση σε πληροφορίες σε κάθε γεωγραφικό σημείο ανεξάρτητα του μέσου, γεγονός που αναδεικνύει κυρίαρχη την ανάγκη για τη διασύνδεση των κατανεμημένων πόρων και υποδομών με ηλεκτρονικά δίκτυα (electronic networks) και ειδικευμένο ενδιάμεσο λογισμικό (middleware) το οποίο θα επιτρέπει την εύκολη και φιλική πρόσβαση των χρηστών αποκρύπτοντας τις ετερογενείς τεχνολογικές υλοποιήσεις των προαναφερθέντων πόρων.

Το κατανεμημένο αυτό περιβάλλον που επιτρέπει το διαμοιρασμό και την από κοινού χρήση υπολογιστικών, αποθηκευτικών και άλλων πόρων, με τη συνδρομή του ενδιάμεσου λογισμικού, ονομάζεται *Πλέγμα Υπολογιστικών Συστημάτων* ή απλά Grid. Η ενοποίηση των δικτύων και του ενδιάμεσου λογισμικού σε μια ενιαία υποδομή με στόχο την

κατανεμημένη αλλά ομογενή πρόσβαση στους πόρους του Grid αναφέρεται ως *Ηλεκτρονική Υποδομή* (e-Infrastructure).

Τα τελευταία χρόνια, η ραγδαία εξάπλωση του διαδικτύου (Internet) σε συνδυασμό με τη διαθεσιμότητα δικτύων κορμού υψηλών ταχυτήτων και τη τεχνολογική ανάπτυξη των υπολογιστών αλλά και του λογισμικού, έχουν δημιουργήσει μια νέα δυναμική στην κλασσική έννοια του όρου «υπολογιστικό περιβάλλον». Στην παρούσα φάση, σημαντικός αριθμός υπολογιστικών πόρων, οι οποίοι περιλαμβάνουν υπολογιστική ισχύ, δεδομένα, υπηρεσίες, εργαλεία λογισμικού, επιστημονικά όργανα, κ.α., βρίσκονται κατανεμημένοι σε παγκόσμιο επίπεδο, δημιουργώντας την ανάγκη για ασφαλή, ομοιόμορφη, αξιόπιστη και απομακρυσμένη πρόσβαση μέσω δικτύων ώστε να αξιοποιηθούν ικανοποιητικά οι δυνατότητες που αυτοί παρέχουν.

Τα Πλέγματα (Grids) είναι μια προσέγγιση της σύστασης δυναμικά δομημένων περιβαλλόντων, χρησιμοποιώντας υπολογιστικούς πόρους που είναι διεσπαρμένοι τόσο γεωγραφικά όσο και οργανωτικά. Ο όρος Grid περιλαμβάνει το σύνολο της υποδομής (υλικό και λογισμικό) και των απαραίτητων υπηρεσιών για τη δημιουργία ενός ενιαίου (γεωγραφικά διεσπαρμένου) υπερ-υπολογιστικού περιβάλλοντος.

Καθώς υπάρχουν πολλοί ορισμοί που όμως αλληλοσυμπληρώνονται και δεν αλληλοαναιρούνται, αναφέρουμε στη συνέχεια τον τεχνικό ορισμό του Grid όπως δίνεται από την IBM [33]:

*«Grid είναι η δυνατότητα, με τη χρήση ενός συνόλου από ανοικτά πρότυπα και πρωτόκολλα, της απόκτησης πρόσβασης σε εφαρμογές, δεδομένα, επεξεργαστική ισχύ, χώρο αποθήκευσης δεδομένων και μίας τεράστιας ποικιλίας από υπολογιστικούς πόρους που διατίθενται στο Internet. Το Grid είναι ένα είδος παράλληλου και κατανεμημένου συστήματος που δίνει τη δυνατότητα να μοιραζόμαστε, να επιλέγουμε και να*

*συγκεντρώνουμε πόρους που κατανέμονται σε διάφορους τομείς διαχείρισης βασιζόμενοι στην διαθεσιμότητα των πόρων τους, την χωρητικότητα, την επίδοση, το κόστος και σε απαιτήσεις Ποιότητας Υπηρεσίας (Quality of Service) που καθορίζονται από το χρήστη.»*

Από τα παραπάνω, γίνεται φανερό ότι ο όρος Grid περιλαμβάνει το σύνολο της υποδομής, υλικό και λογισμικό, κατάλληλα διασυνδεδεμένων μέσω δικτύων υψηλών ταχυτήτων, καθώς και των απαραίτητων υπηρεσιών για τη δημιουργία ενός ενιαίου υπερυπολογιστικού περιβάλλοντος, που αν και είναι γεωγραφικά διεσπαρμένο, εμφανίζεται με τρόπο διαφανή σε όλους τους χρήστες του. Αποτελεί ένα ενιαίο σύνολο υπολογιστικών πόρων, μια συμπαγή - αν και κατανεμημένη - υπολογιστική πλατφόρμα. Το Grid διασυνδέει ετερογενή υπολογιστικά περιβάλλοντα, με όμοια ή διαφορετική φιλοσοφία και υπηρεσίες, δημιουργώντας επιπλέον νέα σύνολα υπηρεσιών με αυξημένες υπολογιστικές δυνατότητες και νέους τρόπους αξιοποίησης των ποικίλων πόρων τους οποίους διαμοιράζει.

### **1.3.2 Υπηρεσιοστρεφείς υποδομές**

Ο όρος υπηρεσιοστρεφής υποδομή (Service Oriented Infrastructure – SOI) [82] αναφέρεται σε μια υπολογιστική υποδομή η οποία ακολουθεί υπηρεσιοστρεφή αρχιτεκτονικό (Service Oriented Architecture – SOA) σχεδιασμό και παρέχεται σαν μια υπηρεσία. Ο όρος είναι αρκετά ευρύς και περιλαμβάνει όλους τους διαμορφώσιμους πόρους της υποδομής όπως υπολογιστικούς, αποθήκευσης, και υλικό και λογισμικό δικτύωσης που απαιτούνται για τη λειτουργία εφαρμογών. Σύμφωνα με τους βασικούς στόχους για τα SOA, οι υπηρεσιοστρεφείς υποδομές διευκολύνουν την επαναχρησιμοποίηση και τη δυναμική κατανομή των απαραίτητων πόρων. Η ανάπτυξη των SOI στρέφεται γύρω από τα χαρακτηριστικά των υπηρεσιών που παρέχονται. Τα χαρακτηριστικά των υπηρεσιών είναι η βάση και για την ανάπτυξη των υποδομών.

Αναλυτικά οι υπηρεσιοστρεφείς υποδομές παρουσιάζονται σε επόμενα κεφάλαια της διατριβής.

### **1.3.3 Ποιότητα υπηρεσιών**

Ο όρος «Ποιότητα Υπηρεσίας» (Quality of Service - QoS) χρησιμοποιήθηκε αρχικά ως ορολογία στα δίκτυα υπολογιστών για να αποδώσει την ικανότητα ενός δικτύου να παρέχει καλύτερες υπηρεσίες για συγκεκριμένη κίνηση (network traffic) χρησιμοποιώντας τα διάφορα πρωτόκολλα δικτύου. Η ανάγκη για αξιόπιστα δίκτυα και η βελτίωση των υπηρεσιών που προσέφεραν αυτά δημιούργησε αυτή την παράμετρο, που τώρα θεωρείται ένας πολύ βασικός συντελεστής αξιολόγησης.

Η σημασία της ποιότητας των παρερχομένων υπηρεσιών σε περιβάλλον υπηρεσιοστρεφών υποδομών και πλεγμάτων είναι πολύ μεγάλη, αφού αυτό που κάνει μια υπηρεσία προσιτή και χρήσιμη στον πελάτη, εκτός από την ίδια την υπηρεσία, είναι η ποιότητά της. Επίσης η ιδιαιτερότητα των πλεγμάτων, πρώτον ως προς τη συνεργασία ετερογενών συστημάτων και δεύτερον ως προς την αρχιτεκτονική τους έκανε πιο εμφανή την ανάγκη προσδιορισμού της ποιότητας των υπηρεσιών. Η συμμετοχή πολλών διαφορετικών συστημάτων στο ίδιο περιβάλλον απαιτεί μια κοινή παράμετρο αξιολόγησής τόσο για τις υποδομές όσο και για τις υπηρεσίες που προσφέρουν. Επιπρόσθετα, η αρχιτεκτονική που χρησιμοποιείται σε τέτοια περιβάλλοντα, βασισμένη στις υπηρεσίες, καθιστά την εισαγωγή του παράγοντα της προσφερόμενης ποιότητας υπηρεσίας εφικτή σε όλα τα επίπεδα της δομής τους.

Τα χαρακτηριστικά της ποιότητας υπηρεσίας μπορούν να διαχωριστούν σε δύο κατηγορίες: αυτά που βασίζονται στα ποσοτικά και αυτά που βασίζονται στα ποιοτικά χαρακτηριστικά μιας υποδομής πλέγματος. Τα πρώτα αναφέρονται σε ιδιότητες όπως η απόδοση του επεξεργαστή, η καθυστέρηση του δικτύου, ο διαθέσιμος αποθηκευτικός

χώρος κλπ., ενώ τα δεύτερα έχουν σχέση με την αξιοπιστία των υπηρεσιών και την ικανοποίηση των απαιτήσεων του χρήστη.

Υπάρχουν αρκετές παράμετροι που εξαιτίας της ιδιομορφίας του περιβάλλοντος πλέγματος και υπηρεσιοστρεφών υποδομών (δυναμική εισαγωγή ετερογενών υπολογιστικών και μη συστημάτων) δεν είναι εύκολο να προσδιοριστούν, καθώς δεν είναι στατικές και εξαρτώνται από την εφαρμογή, ενώ βασικές πληροφορίες για αυτές μπορούν να εξαχθούν από μηχανισμούς παρακολούθησης (monitoring) και άλλα συστήματα/μοντέλα προτυποποίησης των πόρων και των υπηρεσιών.

#### ***1.4 Οργάνωση του εγγράφου***

Το παρόν έγγραφο αποτελείται από έντεκα κεφάλαια. Στις ενότητες των κεφαλαίων αυτών παρουσιάζεται με αναλυτικό τρόπο η ερευνητική εργασία που πραγματοποιήθηκε στα πλαίσια της διατριβής καθώς και τα αποτελέσματα αυτής.

Το **δεύτερο κεφάλαιο** παρέχει γενικές πληροφορίες για τις υπηρεσιοστρεφείς υποδομές και τα περιβάλλοντα πλέγματος, όπως ιστορικά στοιχεία, ορολογία, βασικές αρχές, χαρακτηριστικά, ταξινόμηση και πλεονεκτήματα αυτών στον ευρύτερο τομέα των πληροφοριακών και τηλεπικοινωνιακών συστημάτων.

Στο **τρίτο κεφάλαιο** γίνεται αναφορά σε στοιχεία Αρχιτεκτονικής Σχεδίασης των υπηρεσιοστρεφών υποδομών. Αρχικά αναλύονται οι παράμετροι σχεδίασης και το κεφάλαιο συνεχίζει περιγράφοντας την Αρχιτεκτονική Ανοιχτών Υπηρεσιών Πλέγματος (OGSA) και την Υπηρεσιοστρεφή Αρχιτεκτονική (SOA).

Στο **τέταρτο κεφάλαιο** παρουσιάζεται η έννοια της Ποιότητας Υπηρεσίας (QoS) καθώς και η ανάγκη διασφάλισης αυτής για την υιοθέτηση επιχειρηματικών μοντέλων σε υπηρεσιοστρεφείς υποδομές. Επιπλέον, περιγράφονται οι παράμετροι που καθορίζουν και «ποσοτικοποιούν» την ποιότητα και οι κατηγορίες απαιτήσεων ποιότητας για

συγκεκριμένες παραμέτρους. Το συγκεκριμένο κεφάλαιο καταλήγει αναφέροντας τις αρχιτεκτονικές προκλήσεις που έχουν προκύψει για την διασφάλιση ποιότητας υπηρεσιών σε υπηρεσιοστρεφείς υποδομές.

Στο **πέμπτο κεφάλαιο** περιγράφονται τα πλέγματα επόμενης γενιάς που υλοποιούνται σαν υπηρεσιοστρεφείς υποδομές, οι μηχανισμοί και οι υπηρεσιοστρεφείς αρχιτεκτονικές που τα συνθέτουν, καθώς και τις καινοτομίες και τα πλεονεκτήματα που προκύπτουν από αυτές αναφορικά με την διασφάλιση ποιότητας υπηρεσιών.

Στο **έκτο κεφάλαιο** παρουσιάζει μία καινοτόμο λύση για τη διασφάλιση ποιότητας υπηρεσιών (από την πλευρά του παρόχου) στις υπηρεσιοστρεφείς υποδομές, η οποία αξιοποιεί ιστορικά δεδομένα και έχει δυνατότητες εφαρμογής σε περιβάλλοντα με απαιτήσεις απόκρισης σε πραγματικό χρόνο.

Στο **έβδομο κεφάλαιο** αναλύεται μια ολοκληρωμένη (end to end), καινοτόμος υπηρεσιοστρεφής αρχιτεκτονική προσέγγιση, η οποία εξασφαλίζει ποιότητα υπηρεσιών για διαδραστικές εφαρμογές πραγματικού χρόνου.

Στο **όγδοο κεφάλαιο** αναλύονται οι ανάγκες των εφαρμογών βιοϊατρικής την ένταξη τους σε περιβάλλοντα πλέγματος και υπηρεσιοστρεφείς υποδομές αναφορικά με την ποιότητα υπηρεσιών, ενώ παρουσιάζεται μια νέα αρχιτεκτονική προσέγγιση που δίνει λύση σε ζητήματα υπολογισμού φόρτου εργασίας και χωρητικότητας καθώς και διασφάλισης ποιότητας υπηρεσίας για τις εφαρμογές αυτές.

Τέλος **ένατο κεφάλαιο** περιλαμβάνεται η σύνοψη της διατριβής και τα συμπεράσματα που εξήχθησαν κατά την εκπόνησή της, καθώς και η κύρια συνεισφορά της στον τομέα των υπηρεσιοστρεφών υποδομών και αρχιτεκτονικών. Επιπλέον γίνεται μνεία σε θέματα μελλοντικής εργασίας και επέκτασης των ερευνητικών αποτελεσμάτων.

Τέλος τα **κεφάλαια δέκα και έντεκα** παρουσιάζουν τις συντομογραφίες και βιβλιογραφικές αναφορές αντίστοιχα που χρησιμοποιήθηκαν στη διατριβή.



# 2

## ***Οι υπηρεσιοστρεφείς υποδομές και το περιβάλλον πλέγματος***

Η χρήση ηλεκτρονικών υπολογιστών σε περιβάλλον πλέγματος [37] θεωρείται ολοένα και συχνότερα ως υποδομή νέας γενιάς με δυνατότητα παροχής κατανεμημένων και ετερογενών πόρων με στόχο την παροχή υπολογιστικής ισχύος σε εφαρμογές με υψηλές απαιτήσεις σε πόρους, με διαφανή τρόπο [100]. Στο κεφάλαιο αυτό παρουσιάζονται θέματα για τις τεχνολογίες πλέγματος όπως ιστορικά στοιχεία, όροι, ταξινόμηση και πλεονεκτήματα από τη χρήση αυτών.

### ***2.1 Ιστορικά στοιχεία - Από το διαδίκτυο στο πλέγμα***

Η έρευνα για το Grid ξεκίνησε στους ακαδημαϊκούς χώρους έστω και στα πλαίσια απλών συζητήσεων, πολλά χρόνια πριν ο επιχειρηματικός κόσμος αντιληφθεί τις δυνατότητες που θα του πρόσφερε η μετάβαση σε ένα σύστημα κατανεμημένων υπολογισμών. Όμως οι αρχικές ιδέες, που σχετίζονταν με τον όρο Grid, το αντιμετώπιζαν διαφορετικά απ' ότι οι σημερινές.

Πανεπιστήμια και ερευνητικά ιδρύματα είναι παραδοσιακά αυτοί που εισάγουν νέες ιδέες και προοπτικές στα θέματα που αφορούν τις υπολογιστικές υποδομές. Αυτό ισχύει και στην περίπτωση του διαδικτύου (Internet), πρόιμη μορφή του οποίου αποτελεί το *ARPANET* που αναπτύχθηκε στις ΗΠΑ στη δεκαετία του 1960 από την ερευνητική ομάδα

ARPA (Advanced Research Projects Agency), και το δίκτυο (web) το οποίο αναπτύχθηκε στο ευρωπαϊκό ερευνητικό εργαστήριο του CERN. Στη συνέχεια η επιστημονική κοινότητα ήταν πάλι εκείνη η οποία πρώτη ανακάλυψε τα πλεονεκτήματα των νέων τεχνολογιών καθώς και τρόπου χρησιμοποίησής τους για την επίλυση προβλημάτων.

Υπάρχει πληθώρα τέτοιων προβλημάτων τα οποία μπορούν να επιλυθούν πολύ καλύτερα με τη βοήθεια της προηγμένης τεχνολογίας υπολογιστών. Πολλά από αυτά τα προβλήματα έχουν το κοινό χαρακτηριστικό ότι είναι πολύ απαιτητικά σε υπολογιστική ισχύ. Αυτό σημαίνει ότι όσο πιο μεγάλη υπολογιστική ισχύ διαθέτει κάποιος, τόσο πιο ακριβής είναι η απάντηση στο πρόβλημα του οποίου επιχειρείται η επίλυση. Ισχύει ακόμη ότι ορισμένα προβλήματα δε μπορούν καν να επιλυθούν χωρίς την επαρκή υπολογιστική δύναμη. Το θέμα αυτό των υπολογιστικά απαιτητικών εργασιών αποτελεί συνεχώς τα τελευταία 40 χρόνια της εξέλιξης των υπολογιστικών δομών την κύρια δύναμη που δίνει ώθηση στην εξέλιξη αυτή.

Η ανάγκη για την επίλυση των προβλημάτων αυτών οδήγησε τους ανθρώπους στην ανάπτυξη των υπολογιστών και των υπερυπολογιστών στη συνέχεια. Οι μεγάλοι υπολογιστικοί πόροι της IBM της δεκαετίας του 1960 και οι υπερυπολογιστές Cray στις δεκαετίες 1970 και 1980 κυριαρχούσαν στα υπολογιστικά κέντρα για πάνω από δύο δεκαετίες. Στη συνέχεια επικράτησαν οι αρχιτεκτονικές συμμετρικής πολυεπεξεργασίας SMP και μαζικής παράλληλης επεξεργασίας MPP που περιλάμβαναν πολλούς επεξεργαστές που είχαν τη δυνατότητα να λειτουργούν παράλληλα και για την αποτελεσματικότερη χρήση τους, οι προγραμματιστές εφαρμογών έπρεπε να δομήσουν των κώδικά τους έτσι ώστε να επιτρέπει την παράλληλη εκτέλεσή του. Έπειτα εμφανίστηκαν τα clusters, δηλαδή πολλοί υπολογιστές διασυνδεδεμένοι μεταξύ τους και κατάφεραν να διαδοθούν χάρη στο χαμηλότερο κόστος τους. Έκαναν χρήση προγραμματιστικών μοντέλων όπως το MPI (Message Passing Interface) και το PVM

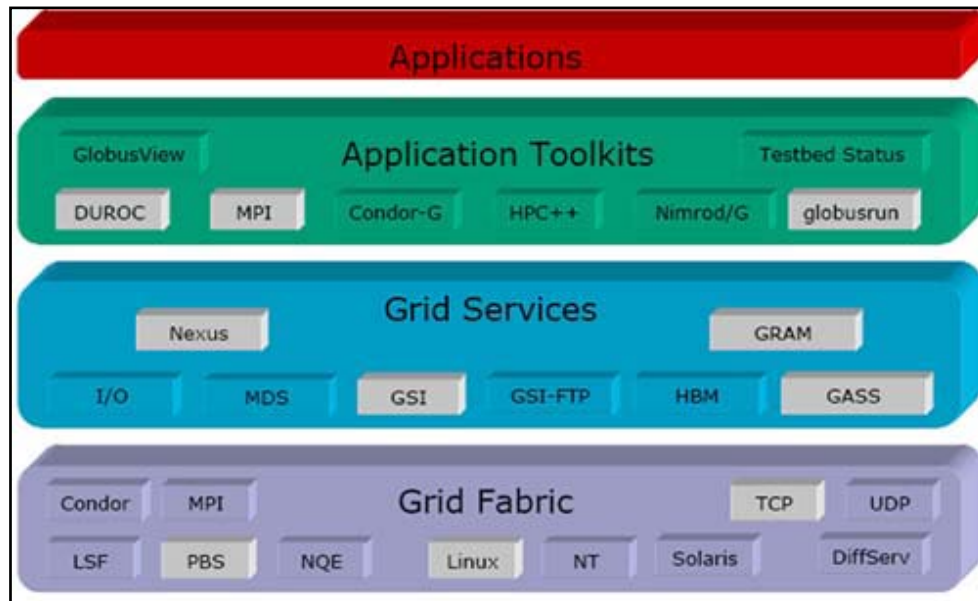
(Parallel Virtual Machine), κατανεμημένου συστήματος αρχείων όπως το NFS (Network File System) και γρήγορων δικτυακών συνδέσεων όπως Ethernet και Myrinet, έτσι ώστε να αποφεύγεται η μείωση της απόδοσης λόγω αργής επικοινωνίας μεταξύ των υπολογιστών που αποτελούν το cluster.

Η αρχιτεκτονική των cluster συστημάτων παρουσίαζε ορισμένα βασικά προβλήματα στην περίπτωση της επικοινωνίας μεταξύ clusters διαφορετικών ιδρυμάτων:

- Τα clusters τυπικά αποτελούνται από ίδιες ή παρόμοιες μηχανές. Δύο clusters διαφορετικών αρχιτεκτονικών ήταν δύσκολο να ενωθούν σε ένα ενιαίο σύστημα. Διαφορές στους compilers, τα εργαλεία, τις βιβλιοθήκες και της δομής των αρχείων δύο διαφορετικών λειτουργικών συστημάτων θα αποτελούσαν προβλήματα για την ενοποίηση.
- Τυπικές παράμετροι για δικτυακές συνδέσεις μεγάλων αποστάσεων απεδείχθησαν μη ικανοποιητικές για την τεχνολογία των clusters. Καθυστέρηση, χαμηλή ρυθμοαπόδοση (throughput), τυχαίες συμφορήσεις του δικτύου (τυπικό για το TCP/IP) και σφάλματα δικτύου ήταν όλοι παράγοντες που συνέβαλαν σε αυτό.
- Δεν υπήρχε η απαιτούμενη τεχνολογία για να αντιμετωπιστούν τα προβλήματα ασφαλείας που προέκυπταν. Συναρτήσεις με χαμηλή ασφάλεια, που ήταν κατάλληλες για εσωτερικά συστήματα που συνήθως προστατεύονταν από firewalls, ήταν τελείως ακατάλληλες για περιβάλλον ανοικτού δικτύου.
- Θέματα διαχείρισης και πολιτικής έδωσαν επίσης μια νέα διάσταση στο πρόβλημα. Τα διάφορα ιδρύματα έπρεπε να καθορίσουν τις συνθήκες κάτω από τις οποίες θα μοιράζονταν τους πόρους τους μέσω του δικτύου. Ακόμη όμως και αν καθορίζονταν μια τέτοια πολιτική, με την υπάρχουσα τεχνολογία δεν υπήρχαν τα απαραίτητα τεχνικά μέσα για τον αποτελεσματικό έλεγχο της και επιβολή της.

Η ερευνητική δραστηριότητα στον τομέα των δικτύων υπολογιστών συνεχίστηκε, με αποτέλεσμα την δημιουργία του *NSFNET* [1986], δικτύου στα 56Kbps που συνέδεε τα πέντε NSF κέντρα υπερ-υπολογιστών. Ως συνέχεια και εξέλιξη αυτών των τεχνολογιών μπορούμε να θεωρήσουμε το πρόγραμμα Condor [1988] του πανεπιστημίου του Wisconsin. Το σύστημα αυτό είναι ένας ‘διαχειριστής φόρτου εργασίας’ (workload manager), με δυνατότητες παρακολούθησης και διαχείρισης πόρων, δρομολόγησης εργασιών και αποτελεί το πρώτο πρόγραμμα με κατεύθυνση προς την αξιοποίηση των Grid υπηρεσιών.

Η ανάπτυξη δικτύων υψηλών ταχυτήτων και η ανάγκη για μεγάλη επεξεργαστική ισχύ οδήγησε σε έντονη ερευνητική δραστηριότητα στον τομέα των Grid τεχνολογιών. Η έρευνα αυτή κατέληξε σε ενδιαφέροντα αποτελέσματα με πιο σημαντικά τα προγράμματα LEGION [1993], SRB [1997], GLOBUS [1998]. Το πρώτο βασίζεται στην ιδέα του ‘εικονικού υπολογιστή (virtual computer): όλοι οι πόροι συνδεδεμένοι μεταξύ τους, εμφανίζονται στον χρήστη ως μία εικονική μηχανή, με αρκετά μειονεκτήματα όμως, όπως η πολύπλοκη υλοποίηση και η μικρή αποδοτικότητα. Το SRB (Storage Resource Broker) ήταν μια πλατφόρμα διαχείρισης αποθηκευτικών πόρων που βοήθησε πολύ στην ανάπτυξη των Grid τεχνολογιών, αφού αντιμετώπισε τα προβλήματα μεταφοράς δεδομένων σε Grid περιβάλλον. Τέλος, το πιο διαδεδομένο σύστημα διαχείρισης Grid υπηρεσιών είναι το GLOBUS, που αναπτύχθηκε στο Argonne National Lab στο πανεπιστήμιο του Berkeley. Το GLOBUS προτυποποίησε πρωτοκόλλα για τη ασφάλεια, μεταφορά δεδομένων, ανακάλυψη πόρων και εκτέλεση εργασιών. Λειτουργεί σε χαμηλό επίπεδο και είναι ανεπτυγμένο σε επίπεδα υπηρεσιών, όπως φαίνεται παρακάτω και στο Σχήμα 3:



Σχήμα 3: Αρχιτεκτονική Globus

Η συνέχεια γίνεται με την ανάπτυξη των Web Services [2001], υπηρεσιών που είναι προσβάσιμες μέσω του διαδικτύου και χρησιμοποιούν συγκεκριμένα πρωτόκολλα περιγραφής (XML, SOAP, WSDL) και τέλος με την εγκαθίδρυση της Αρχιτεκτονικής Ανοιχτών Υπηρεσιών Grid (Open Grid Service Architecture - OGSA) [2002] ως κύριας αρχιτεκτονικής της Grid τεχνολογίας (όπως περιγράφεται στην ενότητα 3.2). Το πρότυπο αυτό είναι το πιο σημαντικό για τις τεχνολογίες του Grid, αφού περιγράφει τις δυνατότητες του συστήματος αυτού να αναλύσει την λειτουργία του σε όλα τα επίπεδά του.

## 2.2 Χαρακτηριστικά

Όπως έχει ήδη αναφερθεί, τα Grids ενοποιούν μέσω ηλεκτρονικών δικτύων υπολογιστικούς, αποθηκευτικούς και άλλους πόρους κατανεμημένους σε τοπική, εθνική και διεθνή κλίμακα. Βάσει αυτού, διακρίνονται με τα εξής χαρακτηριστικά:

- Επιτρέπουν το διαμοιρασμό των πόρων σε πολλαπλούς χρήστες διαφορετικών κοινοτήτων με ετερογενή πεδία εφαρμογών και γεωγραφική κατανομή. Ένα Grid

*Η σελίδα αυτή είναι σκόπιμα λευκή*

μπορεί να στηρίζεται σε ένα τοπικό δίκτυο (campus LAN), μητροπολιτικό δίκτυο MAN, εθνικής εμβέλειας δίκτυο (WAN) ή και διεθνούς κάλυψης δίκτυο όπως το Ευρωπαϊκό Ερευνητικό Δίκτυο GEANT και το Αμερικανικό Abilene ανάλογα με τις απαιτήσεις των εφαρμογών και τις υπάρχουσες δικτυακές υποδομές.

- Απαιτούν ασφαλή πρόσβαση μέσω ενδιάμεσου λογισμικού (middleware) με έμφαση στο λογισμικό ανοικτού κώδικα - open source (π.χ. GLOBUS). Τα Grids επεκτείνουν την φιλοσοφία του ανοικτού λογισμικού σε ανοικτά υπολογιστικά συστήματα, με περιορισμούς μόνο όσο αφορά την ασφάλεια και την διαθεσιμότητα πόρων για την κάλυψη συγκεκριμένων αναγκών.
- Παρουσιάζουν μεγάλη δυνατότητα κλιμάκωσης, με ιδιαίτερα περιορισμένη αρχική επένδυση. Οι αρχιτεκτονικές Grid μπορεί να αποτελέσουν σημαντικό εργαλείο για την υπέρβαση του ψηφιακού χάσματος στον κόσμο, σε μια ήπειρο, σε μία χώρα (κέντρο - περιφέρεια), σε έναν οργανισμό (campus).
- Ενοποιούν μέσω δικτύων Internet / Intranet υπολογιστικές, αποθηκευτικές και άλλες ηλεκτρονικές εγκαταστάσεις με ετερογενείς τεχνολογικές υλοποιήσεις με στόχο την παροχή ολοκληρωμένων Ηλεκτρονικών Υπηρεσιών (eServices). Η ενοποίηση υλοποιείται με χρήση ενός επιπρόσθετου στρώματος ενδιάμεσου λογισμικού (middleware) που αναλαμβάνει το διαμορφασμό των πόρων πάνω από το δίκτυο με τα παραπάνω χαρακτηριστικά.

### **2.3 Οι γενιές των συστημάτων πλέγματος**

Βάσει της ιστορίας και της εξέλιξης των Grids, διακρίνονται οι ακόλουθες γενιές αυτών:

- Σύμφωνα με τον Charlie Catlett, Πρόεδρο του Global Grid Forum – που πλέον ονομάζεται Open Grid Forum [72], η πρώτη γενιά Grids (1st Generation Grids ή *1G Grids*) ουσιαστικά αποτελούνταν από τοπικούς "μετα-υπολογιστές"

(metacomputers) με βασικές λειτουργίες όπως το κατανεμημένο σύστημα αρχείων (sitewide single sign-on), δηλαδή μοναδικό σημείο όπου ο χρήστης δίνει τα προσωπικά στοιχεία του (π.χ. user/password), πάνω στα οποία χτίστηκαν νέες κατανεμημένες εφαρμογές με ειδικά προσαρμοσμένα δικτυακά πρωτόκολλα. Με την υλοποίηση Gigabit test-beds τα 1G Grids επεκτάθηκαν και έγινε προσπάθεια δημιουργίας "Μετα-κέντρων" (metacenters), τα οποία διερεύνησαν θέματα ολοκλήρωσης μεταξύ διαφορετικών κέντρων. Γενικά τα Grids πρώτης γενιάς ήταν εντελώς προσαρμοσμένα στα συγκεκριμένα πειράματα και αποτέλεσαν απόδειξη της ιδέας (proof-of-concept).

- Τα συστήματα 2ης γενιάς Grids, (**2G Grids**), ξεκίνησαν με προγράμματα όπως το Condor, το I-WAY (που αποτέλεσε την αρχή του Globus) και το Legion (που αποτέλεσε την αρχή του Anaki), όπου νέες υπηρεσίες ενδιάμεσου λογισμικού και πρωτοκόλλων επικοινωνιών αποτέλεσαν τη βάση για την ανάπτυξη κατανεμημένων εφαρμογών και υπηρεσιών. Τα Grids 2ης γενιάς ουσιαστικά έδωσαν τα βασικά δομικά στοιχεία, αλλά η χρήση τους απαιτούσε σημαντική προσπάθεια "customization" και διαφόρων εργασιών για να καλυφθούν σημαντικά κενά. Οι ανεξάρτητες αυτές προσπάθειες χρήσης συστημάτων 2ης γενιάς που περιείχαν πολλές "κατά απαίτηση" επεκτάσεις λογισμικού, κατέστησε την διαλειτουργικότητα προβληματική.
- Λαμβάνοντας υπόψη, τόσο την πρότερη εμπειρία από τις 2 πρώτες γενιές, όσο και τις τεχνολογίες των πολύ επιτυχημένων υπηρεσιών web, έχουν ξεκινήσει οι προσπάθειες για την 3η γενιά Grids (**3G Grids**), που βασίζονται στην Αρχιτεκτονική Ανοιχτών Υπηρεσιών Grid (όπως περιγράφεται στην ενότητα 3.2), όπου μια σειρά από προδιαγραφές κοινών και ανοιχτών διεπαφών υποστηρίζουν τη διαλειτουργικότητα ανεξάρτητα ανεπτυγμένων υπηρεσιών. Η πρόσφατα



εκδοθείσα προδιαγραφή Open Grid Services Infrastructure – OGSi [80] είναι ο θεμέλιος λίθος της παραπάνω αρχιτεκτονικής. Με την εισαγωγή προτυποποιημένων τεχνικών προδιαγραφών, η 3η γενιά Grid επιταχύνει τον ανταγωνισμό και την επίτευξη διαλειτουργικότητας όχι μόνο μεταξύ εφαρμογών και εργαλειαθικών, αλλά κυρίως μεταξύ διαφορετικών υλοποιήσεων βασικών υπηρεσιών του Grid.

## **2.4 Ορολογία σχετική με πλέγματα**

### **2.4.1 Το meta-computing**

Το *meta-computing* εμφανίστηκε ως η προσπάθεια για την αποδοτική σύνδεση και συγκέντρωση της υπολογιστικής δύναμης που βρισκόταν σε διάφορα μέρη του κόσμου. Ο όρος προέρχεται από το ότι στα ακαδημαϊκά συστήματα η γνώση για το είδος και την ισχύ των μηχανημάτων αποθηκεύονταν σε έναν κεντρικό meta-υπολογιστή. Οι προσπάθειες συγκεντρώνονταν στην επίλυση υπολογιστικά απαιτητικών προβλημάτων, η οποία διευκολύνονταν από τη συνδυασμένη υπολογιστική ισχύ πολλών επεξεργαστών. Για κάποιο διάστημα το meta-computing θεωρούνταν συνώνυμο του Grid-computing. Όμως τα διάφορα πληροφοριακά συστήματα δε χρησιμοποιούν μόνο επεξεργαστές. Κάθε εργασία έχει είσοδο και έξοδο. Υπάρχουν δεδομένα τα οποία επεξεργάζονται και τα οποία ίσως χρειάζεται να αποθηκευτούν στο τέλος και μάλιστα σε χώρο εκτός του υπολογιστή όπου γίνεται η επεξεργασία. Μπορεί το σύστημα να χρειαστεί να συνεργαστεί με κάποια βάση δεδομένων ή με εξειδικευμένο υλικό (hardware) το οποίο ανήκει σε κάποιο άλλο οργανισμό από αυτόν που θέλει να εκτελέσει την εργασία. Εν τέλει είναι δυνατόν οι πόροι να μην είναι μόνο υπολογιστικοί, αλλά και πόροι λογισμικού, εξειδικευμένο υλικό κλπ. Βλέπουμε λοιπόν ότι το meta-computing αντιμετώπισε μόνο ένα μέρος του προβλήματος.

### **2.4.2 Μη υπολογιστικοί πόροι**

Όπως είπαμε προηγουμένως οι πόροι μπορεί να μην είναι μόνο υπολογιστικοί. Γενικά ως πόρος θεωρείται οποιοδήποτε στοιχείο της δικτυωμένης υποδομής το οποίο διατίθεται προς χρήση μέσω καθορισμένων πρωτοκόλλων Grid. Έτσι ακόμη και τα δίκτυα θεωρούνται ως πόροι οι οποίοι προσφέρουν εύρος ζώνης (bandwidth) για τη μεταφορά δεδομένων. Οι χώροι αποθήκευσης δεδομένων και οι βάσεις δεδομένων θεωρούνται ως πόροι όταν χρησιμοποιούν τυποποιημένες διασυνδέσεις για να επιτρέψουν σε εφαρμογές Grid να διατηρήσουν δεδομένα. Επίσης είναι δυνατόν να υπάρχουν πόροι λογισμικού, γιατί μπορεί να υπάρχουν προγράμματα που βρίσκονται σε συγκεκριμένες τοποθεσίες γιατί μπορούν να τρέξουν μόνο σε εξειδικευμένο υλικό ή τίθεται θέμα του ποιος είναι εξουσιοδοτημένος να τα χρησιμοποιήσει. Κάθε πόρος διαθέτει ορισμένα χαρακτηριστικά τα οποία τον καθιστούν μοναδικό. Τα κυριότερα είναι:

- Απόδοση. Είναι δυνατόν υπολογιστικοί κόμβοι να διαφέρουν ως προς τον αριθμό των επεξεργαστών, των ταχυτήτων τους και τη ποσότητα και ταχύτητα της τοπικής μνήμης. Από την άλλη διάυλοι δικτύου μπορεί να διαφέρουν ως προς το εύρος ζώνης και την καθυστέρηση.
- Αρχιτεκτονική. Για παράδειγμα, αναλόγως της αρχιτεκτονικής ενός επεξεργαστή μπορεί να είναι ή να μην είναι δυνατόν να τρέξει ένα πρόγραμμα σε αυτόν.
- Ποιότητα υπηρεσίας (Quality of Service). Μπορεί για κάποιο χρονικό διάστημα ένα δίκτυο να μην είναι σε θέση να εγγυηθεί την ελάχιστη διαθέσιμη ρυθμοαπόδοση σε ένα χρήστη.
- Αξιοπιστία. Για παράδειγμα, υλικό όπως οι σκληροί δίσκοι χαρακτηρίζονται από το TTF (μέσος χρόνος δραστηριότητας χωρίς σφάλμα).

- Διαθεσιμότητα. Εξαιτίας βλάβης της επικοινωνίας στην περίπτωση των απομακρυσμένων πόρων, η σύνδεση με αυτούς μπορεί να μην είναι δυνατή συνέχεια.
- Δυνατότητες. Για παράδειγμα, κάθε εφαρμογή συνήθως είναι δυνατόν να εκτελεί συγκεκριμένες λειτουργίες και αλγορίθμους, που καμία άλλη εφαρμογή να μη μπορεί.

### **2.4.3 Εικονοποίηση**

Καθώς εξελισσόταν το Grid, κατέστη σαφές ότι θα πρέπει να υπάρχει ένα στρώμα υπηρεσιών (service layer) το οποίο θα επέτρεπε στους χρήστες και τις εφαρμογές τους αιτούνται τη χρήση πόρων χωρίς να γνωρίζουν την ακριβή δομή αυτών. Για παράδειγμα, κάποιος χρήστης θα μπορούσε να ζητήσει μέσω της αντίστοιχης υπηρεσίας την εκτέλεση μιας εργασίας χωρίς να χρειάζεται να ξέρει πως αυτή θα εκτελεστεί (αν θα εκτελεστεί για παράδειγμα από έναν πανίσχυρο επεξεργαστή ή 10 πιο αδύναμους). Το γεγονός αυτό ονομάζεται εικονοποίηση (virtualization) και έχει τα εξής πλεονεκτήματα:

- Καλύτερη κατηγοριοποίηση και παρουσίαση των δυνατοτήτων που βασίζεται περισσότερο στις πραγματικές ανάγκες και όχι σε φυσικούς περιορισμούς (χρειάζομαι 60 επεξεργαστές για μια ώρα, δε με νοιάζει σε πόσα δωμάτια βρίσκονται).
- Πιο αποτελεσματική χρήση των συνηθισμένων πόρων.
- Κοινή πρόσβαση σε σπάνιους πόρους με μοναδικές ικανότητες όπως εξειδικευμένες εφαρμογές και υλικό (hardware).
- Δε χρειάζεται να καθορίζει η εφαρμογή τον τρόπο με τον οποίο η εργασία θα κατανέμεται στους πόρους.

#### **2.4.4 Υπηρεσίες δικτύου**

Η προσπάθεια της επικοινωνίας με απομακρυσμένους πόρους και της χρήσης αυτών, το οποίο αποτελεί βασικό στοιχείο του Grid computing, στηρίχθηκε κατά καιρούς σε διάφορες τεχνολογίες όπως RPC (Remote Procedure Call), CORBA, COM/DCOM, και RMI. Για την επίτευξη όμως αποτελεσματικής επικοινωνίας μεταξύ διαφορετικών συστημάτων χρειάστηκε η εμφάνιση μιας νέας τεχνολογίας με το όνομα 'υπηρεσίες δικτύου (Web Services).

Μία υπηρεσία δικτύου είναι μια καλά καθορισμένη συνάρτηση η οποία είναι προσπελάσιμη μέσω του διαδικτύου. Η επικοινωνία με αυτή γίνεται μέσω της τεχνολογίας XML και του πρωτοκόλλου SOAP που βασίζεται σε αυτή. Η τεχνολογία των υπηρεσιών δικτύου καθορίζει μια κοινή πλατφόρμα επικοινωνίας με την οποία μπορούν να δημιουργηθούν διασυνδέσεις σε διάφορες συναρτήσεις. Η ύπαρξη ενός κοινού συντακτικού (XML) επιτρέπει στις υπηρεσίες να απαλλαγθούν από τις τυπικές ενέργειες που χρειάζονται για την επικοινωνία. Αυτές τις αναλαμβάνουν πλέον οι εκάστοτε εξυπηρετητές που φιλοξενούν τις υπηρεσίες. Έτσι οι υπηρεσίες μπορούν πλέον να γραφτούν εύκολα, να αντικατασταθούν, να αναπτυχθούν και να συντηρηθούν.

Το γεγονός ότι οι πρώτες υπηρεσίες υστερούσαν σε αξιοπιστία, ασφάλεια και απόδοση, όπως επίσης και σε λειτουργικότητα υψηλότερου επιπέδου, είχε ως αποτέλεσμα την εμφάνιση διαφόρων τεχνολογιών που έλυναν τα προβλήματα αυτά. Έτσι οδηγηθήκαμε στην εμφάνιση των WS-Security [13] και του WSRF (WS-Resource Framework) [54]. Το τελευταίο χρησιμοποιείται για τη δημιουργία υπηρεσιών δικτύου οι οποίες προσφέρουν στους μετέχοντες στο Grid πρόσβαση σε πόρους. Μία υπηρεσία δικτύου μπορεί να αντιστοιχεί σε πολλούς πόρους (WS-Resources), ενώ ένα πόρος μπορεί να είναι προσπελάσιμος από διάφορες υπηρεσίες δικτύου. Οι πόροι αυτοί είναι πλέον δυναμικοί, δηλαδή οι ιδιότητές τους μπορούν να αλλάζουν. Μπορούν να δημιουργούνται και να

καταστρέφονται κατ' απαίτηση και η πρόσβαση σε αυτούς είναι δυνατή μόνο μέσω μιας δικτυακής υπηρεσίας.

#### **2.4.5 Εικονικοί οργανισμοί**

Το ακριβές πρόβλημα το οποίο υποκινεί την ανάπτυξη του Grid είναι ο ελεγχόμενος και συντονισμένος διαμοιρασμός και χρήση πόρων για την επίλυση προβλημάτων στο πλαίσιο δυναμικών Εικονικών Οργανισμών, E.O. (Virtual Organizations-VOs) [38]

**Error! Reference source not found.** Ο παραπάνω διαμοιρασμός αφορά όχι μόνο στην ανταλλαγή δεδομένων αλλά επίσης στην άμεση πρόσβαση σε οντότητες Grid (υπολογιστικές μονάδες, υπηρεσίες, λογισμικό, δεδομένα και άλλους πόρους), που συνδέονται μεταξύ τους σύμφωνα με ένα πλαίσιο εμπιστοσύνης. Επιπλέον, ο προαναφερθείς διαμοιρασμός θα πρέπει να είναι ελεγχόμενος, με τους παρόχους και τους χρήστες των πόρων να ακολουθούν πρωτόκολλα τα οποία θα καθορίζουν με σαφήνεια τι θα πρέπει να μοιραστεί, ποιος επιτρέπεται να διαμοιράσει και ποιες είναι οι συνθήκες κάτω από τις οποίες πραγματοποιείται ο διαμοιρασμός. Τα μέλη ενός E.O. επιδιώκουν την επίτευξη ενός κοινού στόχου και λόγω αυτού οι διάφορες οντότητες μπορούν να γίνονται μέλη ή να αποχωρήσουν από έναν E.O. δυναμικά και ενώ το σύστημα βρίσκεται σε λειτουργία. Για παράδειγμα ένα πρόγραμμα εκτελεί μια συνάρτηση, η οποία για την εκτέλεσή της χρειάζεται πρόσβαση σε διάφορες υπηρεσίες, που βρίσκονται σε διάφορα μέρη. Είναι δυνατόν στην περίπτωση αυτή να δημιουργηθεί δυναμικά για όσο διάστημα χρειαστεί ένας E.O., ο οποίος θα καθορίζει το επίπεδο ασφάλειας μεταξύ των μελών του, παρέχοντας (κάτω από προϋποθέσεις βέβαια) στη συνάρτηση, πρόσβαση στις αναγκαίες υπηρεσίες.

Επομένως ένας ορισμός του E.O. μπορεί να είναι: το σύνολο των οντοτήτων που συνδέονται προσωρινά για ένα κοινό σκοπό ή για την εκτέλεση μιας κοινής εργασίας. Ένα Grid μπορεί να περιλαμβάνει πολλούς E.O., ενώ μια οντότητα μπορεί να είναι μέλος

πολλών Ε.Ο. ταυτόχρονα με αποτέλεσμα να υπάρχει δυνατότητα οι VO να επικαλύπτονται μεταξύ τους.

Βάσει των παραπάνω, υπάρχουν πολλές απαιτήσεις για την επιθυμητή λειτουργία των Ε.Ο., όπως :

- Ευέλικτες σχέσεις διαμοιρασμού: Οι σχέσεις διαμοιρασμού μπορούν να μεταβάλλονται δυναμικά στον χρόνο, εννοώντας το ποιοι πόροι διαμοιράζονται, τον τύπο την πρόσβασης που επιτρέπεται, και σε ποιους συμμετέχοντες επιτρέπεται η χρήση του πόρου. Αυτές οι σχέσεις δεν ονομάζουν απαραίτητα ένα σύνολο συμμετεχόντων πόρων, αλλά τις περισσότερες φορές καθορίζονται εμμέσως από τις πολιτικές πρόσβασης του πόρου. Για παράδειγμα ένας οργανισμός μπορεί να επιτρέψει την πρόσβαση σε ένα πόρο σε οποιονδήποτε μπορεί να αποδείξει ότι είναι πελάτης.
- Μηχανισμοί εντοπισμού: Η δυναμική φύση των σχέσεων διαμοιρασμού απαιτεί την ύπαρξη μηχανισμών για την ανακάλυψη και τον προσδιορισμό των σχέσεων που υπάρχουν σε μια συγκεκριμένη χρονική στιγμή. Για παράδειγμα ένας νέος συμμετέχων σε μια Ε.Ο. πρέπει να έχει την δυνατότητα να εντοπίσει τους πόρους στους οποίους του επιτρέπεται η πρόσβαση, την ποιότητα υπηρεσίας (QoS) που παρέχουν αυτοί οι πόροι κτλ.
- Επίλυση θεμάτων χρονοπρογραμματισμού και ταυτόχρονης δέσμευσης ενός συνόλου πόρων: Πολλά προβλήματα απαιτούν την ταυτόχρονη χρήση πόρων για να επιλυθούν. Γι' αυτόν τον λόγο οι σχέσεις διαμοιρασμού πρέπει να μπορούν να συνδυαστούν για την συντονισμένη χρήση πολλών πόρων που βρίσκονται σε κάποιους οργανισμούς. Για παράδειγμα μια διαμοιραζόμενη υπολογιστική μονάδα (πόρος 1) που επεξεργάζεται δεδομένα από μια διαμοιραζόμενη μονάδα αποθήκευσης (πόρος 2).

- Μηχανισμοί Αποστολής Δικαιωμάτων (delegation): Μηχανισμοί με τους οποίους ο χρήστης εξουσιοδοτεί τον διαμοιραζόμενο πόρο με τα δικαιώματά του. Για παράδειγμα μια υπολογιστική μονάδα να εξουσιοδοτηθεί ώστε να έχει δικαίωμα πρόσβασης στα αρχεία της οντότητας που την χρησιμοποιεί.
- Σχέσεις ομότιμων: Οι σχέσεις διαμοιρασμού πολύ συχνά δεν είναι απλά της μορφής πελάτης - εξυπηρετητής αλλά ομότιμων οντοτήτων, δηλαδή οι παροχές μπορεί να είναι και καταναλωτές.
- Διαφορετική λειτουργία ενός πόρου: Ο ίδιος πόρος μπορεί να χρησιμοποιηθεί με διαφορετικό τρόπο ανάλογα με τους περιορισμούς πρόσβασης και τον σκοπό για τον οποίο γίνεται η διαμοίραση. Για παράδειγμα ένας υπολογιστής σε μια Ε.Ο. μπορεί να χρησιμοποιείται για την εκτέλεση ενός συγκεκριμένου λογισμικού και σε κάποια άλλη Ε.Ο. να χρησιμοποιείται γενικά, παρέχοντας υπολογιστική ισχύ.

## **2.5 Ταξινόμηση των πλεγμάτων**

Υπάρχουν διάφορες ταξινομήσεις συστημάτων πλέγματος σε επίπεδο εφαρμογών, επαγγελματικής αξίας, επιστημονικής αποτελεσματικότητας και αρχιτεκτονικής. Μια από τις δημοφιλέστερες ταξινομήσεις σχετίζεται με το επίπεδο πολυπλοκότητας της οργάνωσης και της έκτασης που καταλαμβάνει το όλο σύστημα. Βάσει αυτής διακρίνονται οι ακόλουθες κατηγορίες:

- Departmental Grids, τα οποία εγκαθίστανται συνήθως σε ένα τομέα ενός οργανισμού και συνήθως προορίζονται για να εξυπηρετήσουν ένα και μόνο σκοπό. Εξωτερικά προστατεύονται με firewall κάτι το οποίο ταιριάζει στην περίπτωση αφού συνήθως η όποια επικοινωνία γίνεται εσωτερικά στον τομέα. Η περίπτωση αυτή είναι παρόμοια με αυτή των clusters. Υπάρχει όμως η διαφορά ότι το hardware που χρησιμοποιείται δε παρουσιάζει απαραίτητα ομοιογένεια. Επίσης

υπάρχει μικρή μόνο ανάγκη για τον καθορισμό πολιτικής πρόσβασης και για την επίβλεψη της πρόσβασης. Τελικά ο σκοπός των departmental Grids είναι να συγκεντρώσουν και εικονοποιήσουν τους πόρους του τομέα έτσι ώστε να επιτευχθεί η καλύτερη δυνατή χρήση τους.

- Enterprise Grids, τα οποία ανήκουν σε μία οργανωμένη οντότητα όπως μια επιχείρηση και χρησιμοποιούνται για πολλούς σκοπούς. Υπάρχει η δυνατότητα επικοινωνίας με τον εξωτερικό κόσμο, η οποία δεν είναι αμφίδρομη καθώς οι πόροι καθαυτοί δε μετακινούνται έξω από τα όρια του firewall. Τα enterprise grids χρησιμοποιούνται από οργανισμούς που θέλουν με κάποιο τρόπο να συγκεντρώσουν τους πόρους τους που πριν ήταν διασκορπισμένοι σε διάφορα τμήματα του Φ. Στην περίπτωση αίτησης χρήσης πόρων από δύο ή και περισσότερους τομείς της επιχείρησης εφαρμόζονται μηχανισμοί επίλυσης τέτοιων διενέξεων μέσω ενός συστήματος πολιτικών χρήσης (usage policies).
- Partner Grids, τα οποία είναι εγκαταστάσεις που ξεπερνούν το firewall μιας επιχείρησης και επιτρέπουν τη κοινή χρήση των πόρων από διάφορους οργανισμούς. Η ανάγκη για τα partner grids προέκυψε από την ανάγκη συνεργασίας μεταξύ διάφορων επιχειρησιακών οντοτήτων για την επίτευξη ενός κοινού στόχου. Η έννοια των Partner Grids δε θα πρέπει να συγχέεται με αυτή των Εικονικών Οργανισμών (όπως αναφέρθηκε στην παράγραφο 2.4.4). Ένας E.O. μπορεί να περιλαμβάνει πολλά partner grids και ένα partner grid πολλούς E.O.. Η διάθεση πόρων στο εξωτερικό περιβάλλον ενός οργανισμού απαιτεί την ύπαρξη προτύπων και πολιτικών για την ασφάλεια, την επίβλεψη των διάφορων ενεργειών και το accounting.
- Open Grids, τα οποία αναμένεται να εμφανιστούν στο μέλλον και αναφέρονται σε μια πλατφόρμα που αποτελείται από υποδομή, μεσολογισμικό (middleware) και



εφαρμογές που χρησιμοποιούνται από κοινού από διάφορους ανεξάρτητους οργανισμούς. Υπό φυσιολογικές συνθήκες ένα open grid δεν είναι αφιερωμένο στην επίτευξη ενός και μόνο στόχου. Αντίθετα οι συμμετέχοντες χρησιμοποιούν το σύστημα για να επιτύχουν τους δικούς τους ανεξάρτητους στόχους, άλλες φορές δουλεύοντας μεμονωμένα και άλλες φορές συμμετέχοντας σε εικονικούς οργανισμούς. Το grid θα προσφέρει τους πόρους και τις απαραίτητες υπηρεσίες για το έργο, καθώς και ένα μέσο για το sharing και την επικοινωνία μεταξύ των συνεργαζομένων. Οντότητες που δε χρειάζονται πλέον τους πόρους τους, θα μπορούν να τους αξιοποιούν διαθέτοντάς τους σε ένα από τα υπάρχοντα grid και αποκτώντας σε αντάλλαγμα οικονομικό κέρδος. Το Open Grid (συχνά αναφέρεται και ως παγκόσμιο Grid) θα αποτελέσει φυσικό επακόλουθο της παράλληλης ύπαρξης διάφορων enterprise grids και partner grids που θα επικοινωνούν μεταξύ τους με τη χρήση των ίδιων πρωτοκόλλων.

## **2.6 Αρχιτεκτονική σχεδίαση του πλέγματος**

Η διαδικασία της δημιουργίας μιας αρχιτεκτονικής Grid είναι παρόμοια με άλλες προσπάθειες κατασκευής ενός συστήματος, ειδικά όταν αναφερόμαστε σε κατανεμημένα συστήματα. Παρόλα αυτά, ορισμένοι παράμετροι διαδραματίζουν σημαντικότερο ρόλο στο σχεδιασμό ενός Grid συστήματος απ' ότι άλλοι. Αυτοί οι παράμετροι είναι:

- Απαιτήσεις ασφάλειας
- Ευαισθησία δεδομένων
- Μέγιστες απαιτήσεις σε επεξεργαστική ισχύ
- Αποθήκευση δεδομένων
- Διαθέσιμο εύρος ζώνης στο διαδίκτυο
- Υπάρχοντες πόροι

- Πόροι ειδικού σκοπού
- Μεταφερσιμότητα
- Δυνατότητες συνεργασίας

### **2.6.1 Απαιτήσεις ασφάλειας**

Τυπικά δεδομένα επιχειρήσεων, επιχειρησιακές πρακτικές και επιχειρησιακοί πόροι πρέπει να προστατεύονται από καταστροφή, κλοπή και γενικά από ενέργειες που στοχεύουν να βλάψουν το σύστημα. Σχεδόν όλες οι σύγχρονες εταιρείες σχεδιάζουν κανόνες ασφάλειας, έτσι ώστε να επιτρέπουν στους υπαλλήλους τους να έχουν πρόσβαση στα δεδομένα της εταιρείας και ταυτόχρονα να αποτρέπουν την πρόσβαση από οποιονδήποτε θέλει να βλάψει τα συμφέροντά της.

Κατά καιρούς έχουν εμφανιστεί διάφορες λύσεις. Τα firewalls ήταν μια γρήγορη και εύκολη προσέγγιση, η οποία όμως υστερούσε σε ευελιξία. Τα firewalls επιτρέπουν μόνο τη κυκλοφορία δεδομένων, η οποία έχει εγκριθεί από το διαχειριστή του δικτύου. Έτσι κάθε δραστηριότητα για την οποία δεν έχει δοθεί ειδική έγκριση θεωρείται επίθεση προς το σύστημα. Ενώ όμως το σύστημα προστατεύεται, δημιουργείται πρόβλημα απόδοσης σε επικοινωνίες διαφόρων ειδών, αφού πριν γίνει κάθε επικοινωνία θα πρέπει να υπάρξει συνεννόηση με το διαχειριστή του δικτύου, έτσι ώστε να ρυθμίσει το firewall (ανοίγοντας ίσως και κάποιες πόρτες) και τελικά να επιτραπεί η επικοινωνία..

Με τη μετάβαση της υποδομής του Grid στις τεχνολογίες των υπηρεσιών δικτύου (Web Services) και XML, καθίσταται ευκολότερο για ένα firewall να επιβλέπει τα δεδομένα που περνούν. Επίσης με την εμφάνιση των WS-Agreements [1] (ή SLA-Service Level Agreement), επιτρέπεται η διαπραγμάτευση για την πραγματοποίηση της επικοινωνίας ή μη, να γίνεται με παρόμοιο τρόπο για όλων των ειδών τις υπηρεσίες. Σε αυτή την περίπτωση μπορεί πλέον το firewall να διατηρεί ανοικτές διασυνδέσεις, με τις οποίες θα

συνδέονται οι διάφορες εφαρμογές, για να πραγματοποιήσουν τη διαπραγμάτευση με το σύστημα, αποκτώντας τελικά εξουσιοδοτημένη πρόσβαση σε υπηρεσίες και δεδομένα.

### **2.6.2 Ευαισθησία δεδομένων**

Στην περίπτωση που τα διαχειριζόμενα δεδομένα είναι ευαίσθητα, το περιβάλλον του Grid είναι πολύ δύσκολο να παρέχει αρκετή ασφάλεια. Για παράδειγμα, είναι πολύ εύκολο για κάποιον να παρεμβληθεί στην επικοινωνία και να αλλάξει τα μεταφερόμενα δεδομένα. Στην περίπτωση όμως των σχετικά κλειστών Grid, όπως τα enterprise Grids, η ασφάλεια των δεδομένων είναι σχετικά εύκολο να εξασφαλιστεί. Σε πιο ανοιχτές αρχιτεκτονικές όπως τα partner Grids υπάρχουν τρία (3) επίπεδα στα οποία πρέπει να εξασφαλιστεί ασφάλεια των δεδομένων:

- Μια εργασία μπορεί να πραγματοποιηθεί σε οποιονδήποτε εξουσιοδοτημένο κόμβο του Grid. Σε αυτή την περίπτωση μια λύση είναι η δημιουργία εικονικών οργανισμών που θα καθορίζουν ποιοι κόμβοι είναι αξιόπιστοι.
- Ο χώρος στον οποίο αποθηκεύονται τα δεδομένα ή κρατούνται αντίγραφά τους. Ο χώρος που αποθηκεύονται τα αυθεντικά δεδομένα είναι εύκολο να ασφαλιστεί, αφού συνήθως ανήκει στον οργανισμό τον οποίο τα δεδομένα αφορούν περισσότερο. Στη περίπτωση των χώρων που κρατούνται αντίγραφα, μπορεί πάλι να δημιουργηθεί εικονικός οργανισμός που θα καθορίζει σε ποιους κόμβους επιτρέπεται να η αποθήκευση των δεδομένων.
- Τέλος σημαντικό είναι να υπάρχει ασφάλεια των δεδομένων κατά τη μεταφορά τους. Σήμερα ο καθορισμός ασφαλών καναλιών μεταφοράς δεδομένων είναι δυνατός με τη χρήση των προτύπων SSL (Secure Socket Layer) [5] και WSS (Web Services Security) [41]. Έτσι είναι δυνατή η κρυπτογράφηση των δεδομένων που μεταφέρονται μεταξύ κόμβων.

### **2.6.3 Μέγιστες απαιτήσεις σε επεξεργαστική ισχύ**

Πριν την εμφάνιση του Grid, όταν μια επιχείρηση ήθελε να πραγματοποιήσει μια εργασία αγόραζε επεξεργαστική ισχύ με τη μορφή επεξεργαστών που εγκαθίσταντο σε υπολογιστές ή υπερυπολογιστές. Η ισχύς αυτή πολλές φορές ξεπερνούσε την απαιτούμενη, αφού οι απαιτήσεις για επεξεργαστική ισχύ συγκεκριμένες ώρες της ημέρας μπορούσαν να είναι πολύ μικρότερες. Επίσης μπορεί να απευθυνόταν μόνο στις ανάγκες της συγκεκριμένης εργασίας και να μην ήταν δυνατό να χρησιμοποιηθεί και για άλλες εφαρμογές της επιχείρησης. Έτσι η πλεονάζουσα ισχύς δεν αξιοποιούνταν.

Με την εμφάνιση όμως του Grid όμως κατέστη δυνατό η επιχείρηση να χρησιμοποιεί την ισχύ για περισσότερες εφαρμογές, όπως επίσης να διαθέτει επεξεργαστική ισχύ σε τρίτους αποκομίζοντας και κάποιο κέρδος. Επιπλέον, η επιχείρηση έχει τη δυνατότητα να μην αγοράσει επεξεργαστική ισχύ που θα ανταποκρίνεται στις μέγιστες απαιτήσεις της, εφόσον βέβαια αυτές οι απαιτήσεις δεν υφίστανται για μεγάλο διάστημα της ημέρας. Αντίθετα έχει τη δυνατότητα, για τις ώρες που οι απαιτήσεις φτάνουν στο maximum, να καταφεύγει σε προσφορές επεξεργαστικής ισχύς από τρίτους για να καλύψει τις βραχυπρόθεσμες απαιτήσεις της.

### **2.6.4 Αποθήκευση δεδομένων**

Το Grid θεωρείται ως επίπεδη αρχιτεκτονική από την άποψη ότι όλοι οι κόμβοι είναι ισοδύναμοι. Η αποθήκευση των δεδομένων ακολουθεί την ίδια λογική και επομένως οποιοδήποτε δεδομένο μπορεί να τοποθετηθεί οπουδήποτε. Το τι θα αποθηκευτεί σε μια συγκεκριμένη τοποθεσία καθορίζεται από το σύστημα αποθήκευσης, τη διαθέσιμη χωρητικότητα, το ρυθμό δεδομένων, το κόστος αποθήκευσης, το χρόνο ζωής των δεδομένων και φυσικά την πολιτική.

Συνήθως τα πρωτόκολλα που χρησιμοποιούνται εδώ και που διαχειρίζονται την αποθήκευση των δεδομένων σύμφωνα με την απαίτηση του χρήστη λειτουργούν με

σχετικά αδιαφανή τρόπο. Αυτό σημαίνει ότι ο χρήστης δεν γνωρίζει το μηχανισμό, με τον οποίο αποθηκεύεται ένα αρχείο του (αν πχ. αποθηκεύεται σε δίσκο ή ταινία). Αυτό που γνωρίζει είναι ότι έχει κάνει μια συμφωνία με το πρωτόκολλο, η οποία του υπόσχεται ότι το αρχείο θα αποθηκευτεί και θα διατηρηθεί στο χώρο αποθήκευσης. Το σημαντικό στην περίπτωση του Grid είναι ότι με τη χρήση των κατάλληλων πρωτοκόλλων και προτύπων τα διάφορα συστήματα αποθήκευσης αντιμετωπίζονται με ενιαίο τρόπο.

### **2.6.5 Διαθέσιμο εύρος ζώνης στο διαδίκτυο**

Αφού το Grid στηρίζεται στο διαδίκτυο, το διαθέσιμο εύρος ζώνης (bandwidth) αποτελεί σημαντικό παράγοντα. Επομένως το μέγεθος της σύνδεσης με το internet μιας εφαρμογής που τρέχει πάνω σε Grid, είναι σημαντικό. Όταν κάποιος αγοράζει πρόσβαση σε πόρους, πρέπει να υπάρχει αρκετό εύρος ζώνης, για να προωθηθεί η εργασία μέσω του διαδικτύου. Μια εφαρμογή που τρέχει πάνω σε Grid θα πρέπει να μεταφέρει μαζί της το περιβάλλον λειτουργίας της, γεγονός που στη περίπτωση που τρέχουν πολλές μικρές εφαρμογές μαζί, καθιστά το μέγεθος των μεταφερόμενων δεδομένων (εφαρμογή, βιβλιοθήκες, αρχεία) αρκετά μεγάλο. Μάλιστα, στην περίπτωση που τα αποτελέσματα μιας έστω και μικρής εφαρμογής παράγουν μεγάλο όγκο αποτελεσμάτων, τα πράγματα χειροτερεύουν.

Σήμερα, στην περίπτωση τουλάχιστον των επιστημονικών εφαρμογών, οι συνδέσεις με το διαδίκτυο μπορεί να είναι και της τάξης των μερικών δεκάδων Gigabit/sec. Παρόλα αυτά στη καθόλου ασυνήθιστη περίπτωση που 100 κόμβοι επιχειρούν ταυτόχρονα να εκτελέσουν μια εργασία με σχετικά μεγάλο όγκο δεδομένων, ακόμη και το εύρος ζώνης μερικών Gbps μπορεί να μην αποδειχθεί αρκετό καθυστερώντας έτσι εν τέλει την εκτέλεση των εργασιών. Επειδή είναι χαρακτηριστικό του Grid να υπάρχουν μεγάλοι όγκοι δεδομένων και χρήστες που τους διαχειρίζονται και είναι κατανεμημένοι γεωγραφικά, πρέπει να λαμβάνονται υπόψη τα παρακάτω:

- Σχεδιασμός μιας αρχιτεκτονικής δεδομένων που να εξυπηρετεί τις ανάγκες. Θα πρέπει να λαμβάνονται υπόψη τα είδη των συνδέσεων των διαφόρων χρηστών, που βρίσκονται, ποιος είναι ο αναμενόμενος όγκος δεδομένων και ποιο το απαιτούμενο εύρος ζώνης και ποιότητα υπηρεσίας (QoS).
- Θα πρέπει συχνά να γίνεται προσωρινή αποθήκευση επιπλέον αντιγράφων των δεδομένων (caching), για να βελτιστοποιείται ο χρόνος πρόσβασης. Ερωτήσεις που θα πρέπει να απαντηθούν είναι αν το αντίγραφο είναι ενημερωμένο, πόσο συχνά να ανανεώνεται και στη περίπτωση του κατακευματισμένου συστήματος caching, ποιο αντίγραφο είναι πλησιέστερο στο χρήστη.
- Για αξιοπιστία μπορεί να χρησιμοποιηθεί πλεονασμός δεδομένων έτσι ώστε να εξασφαλίζεται ότι τα δεδομένα είναι διαρκώς διαθέσιμα. Έτσι αν ένας κόμβος που έχει το δεδομένο πάψει να λειτουργεί, τότε αυτό μπορεί να αναζητηθεί αυτόματα σε άλλον.
- Για τη μεταφορά των δεδομένων μεταξύ των τόπων αποθήκευσης και χρήσης τους, χρειάζεται ένα μέσο, το οποίο να παρέχει γρήγορη, αξιόπιστη και ασφαλή μεταφορά.
- Επίσης πρέπει να καθορίζεται σε ποιον ανήκουν τα δεδομένα, ποιοι μπορούν να τα δουν και ποιοι να τα τροποποιήσουν ή να τα σβήσουν.
- Σε ορισμένες περιπτώσεις χρειάζονται επιπλέον δεδομένα μεταπληροφορίας που θα περιγράφουν τη δομή του συστήματος και των δεδομένων αυτού. Αυτά τα δεδομένα έχουν παρόμοιες ανάγκες με τα υπόλοιπα, όπως να είναι προσβάσιμα και αξιόπιστα.

### **2.6.6 Υπάρχοντες πόροι**

Η δυνατότητα της επαναχρησιμοποίησης υπαρχόντων πόρων, συνήθως υπολογιστών, ήταν από τους κύριους λόγους που οδήγησαν στην αποδοχή της τεχνολογίας του Grid.

Σημαντικό για ένα Grid είναι η δημιουργία ενός είδους ευρετηρίου πόρων καθώς και της ισχύς τους.

Υπάρχουν μερικές κατηγορίες τέτοιων υπολογιστικών πόρων:

- Υπολογιστές γραφείου (desktop PCs), οι οποίοι υπάρχουν σχεδόν σε όλο τον κόσμο και όχι μόνο σε επιχειρήσεις, αλλά και σε απλά νοικοκυριά. Συνήθως αυτοί οι πόροι είναι διαθέσιμοι τις νυχτερινές ώρες και χαρακτηρίζονται από υψηλό κόστος συντήρησης ανά μηχανήμα. Ακόμη και έτσι όμως το πλήθος τους, σε συνδυασμό με το ότι παραμένουν ανενεργοί για 12 με 16 ώρες τη μέρα, τους καθιστά ένα ισχυρό εργαλείο για εργασίες (συνήθως συμπληρωματικές) που γίνονται νυχτερινές ώρες.
- Clusters. Αυτά τα συστήματα είναι συνήθως διαθέσιμα 24 ώρες το εικοσιτετράωρο και είναι αξιόπιστα και ασφαλή. Η γνώση της δομής ενός cluster, τι μηχανήματα περιλαμβάνει και ποιους εξειδικευμένους πόρους, είναι πολύτιμη για ένα Grid σύστημα. Έτσι με τη χρήση του κατάλληλου συστήματος, (scheduler) μπορούν να αντιστοιχιστούν οι εργασίες στους πόρους που είναι καταλληλότεροι για την εκτέλεσή τους.
- Συστήματα συμμετρικής πολυεπεξεργασίας (SMP). Αυτά τα συστήματα μέχρι την εμφάνιση της Grid τεχνολογίας παρέμεναν αναξιοποίητα για μεγάλα χρονικά διαστήματα. Πλέον είναι δυνατή η χρήση τους ως πόροι και είναι ιδιαίτερα χρήσιμα στην περίπτωση εργασιών που εκμεταλλεύονται ακριβώς τις ειδικές δυνατότητες που διαθέτουν.

Σημαντικός είναι ο καθορισμός ποια είδη εργασιών εκμεταλλεύονται καλύτερα τις ειδικές ικανότητες κάθε συστήματος. Όταν δημιουργηθεί ένα ευρετήριο των πόρων που είναι διαθέσιμοι, της ζήτησης που μπορεί να καλυφθεί με τη διάθεση αυτών στο Grid και των προτεραιοτήτων των εργασιών, μπορεί κάποιος να διαπιστώσει, εάν η υπάρχουσα

υποδομή επαρκεί για να καλύψει τις ανάγκες του ή αν χρειάζεται να ξοδευτούν επιπλέον χρήματα για την ενίσχυσή της.

### **2.6.7 Πόροι ειδικού σκοπού**

Πολλοί οργανισμοί διαθέτουν πόρους οι οποίοι δημιουργήθηκαν ειδικά για την εκτέλεση μιας εργασίας ή για μια κατηγορία από εργασίες. Αυτοί οι πόροι μπορεί να είναι ειδικά μηχανήματα, όπως συσκευές μετρήσεων, ειδικού σκοπού λογισμικό και ειδικού σκοπού υλικό. Παραδοσιακά αυτοί οι πόροι θα γίνονταν προσβάσιμοι μέσω ειδικών πρωτοκόλλων ή ειδικά διαμορφωμένων δικτύων.

Με την έλευση του Grid computing εμφανίστηκαν διάφορα εργαλεία, που επιτρέπουν τη δημιουργία νέων υπηρεσιών κατάλληλων για σύνδεση με αυτούς τους πόρους. Βέβαια τα εργαλεία αυτά δεν έχουν φτάσει στο επίπεδο ωριμότητας που απαιτείται, ώστε να είναι δυνατό αυτό για όλους τους πόρους. Για τους ειδικούς πόρους που αυτό δεν είναι δυνατό μέχρι στιγμής, υπάρχουν δύο λύσεις:

- Χάρη στις τεχνολογίες των υπηρεσιών δικτύου και XML υπάρχει διαθέσιμη μια τεράστια ποικιλία βιβλιοθηκών, που επιτρέπουν τη δημιουργία κώδικα, που θα ανταποκρίνεται στις εκάστοτε ανάγκες. Ο κώδικας αυτό μπορεί να αποτελεί ένα resource wrapper ή minihosting περιβάλλον, επιτρέποντας την πρόσβαση στον πόρο μέσω των υπάρχοντων ιδιόκτητων διασυνδέσεων.
- Μια άλλη επιλογή είναι να τρέχει σε ένα μηχάνημα μια υπηρεσία Grid, με την οποία θα επικοινωνεί αυτός που επιθυμεί την πρόσβαση στον πόρο. Η Grid υπηρεσία αυτή θα προωθεί μέσω ειδικού μηχανισμού τις απαραίτητες πληροφορίες στον πόρο. Η δεύτερη αυτή λύση προβλέπεται να είναι φθηνότερη και μάλλον θα προτιμηθεί στο μέλλον.



### **2.6.8 Δυνατότητες συνεργασίας**

Ένα από τα σημαντικότερα ζητήματα κατά το σχεδιασμό ενός Grid συστήματος, είναι οι δυνατότητες που αυτό θα προσφέρει στους χρήστες του για συνεργασία με άλλους χρήστες και οντότητες του συστήματος. Εξάλλου βασικός στόχος του Grid είναι η επίτευξη κοινών στόχων, όπως η επίλυση ενός πολύπλοκου προβλήματος, με τη συνεργασία πολλών παραγόντων, πόρων κλπ.

Σημαντικό επομένως είναι να υπάρχει η δυνατότητα καθορισμού πολιτικών διαπραγμάτευσης για την επίτευξη της συνεργασίας. Πρέπει να καθορίζονται πρωτόκολλα συνεργασίας και κανονισμοί, οι οποίοι θα εξασφαλίζουν την ασφάλεια της όλης διαδικασίας.

Πρέπει να δίνεται η δυνατότητα δημιουργίας δυναμικών εικονικών οργανισμών και δυναμικής διαχείρισής τους. Η έννοια του εικονικού οργανισμού είναι καθοριστική, καθώς σύμφωνα με τον ορισμό του, οι μετέχοντες σε έναν εικονικό οργανισμό συνεργάζονται μεταξύ τους για να εξυπηρετήσουν τα κοινά τους συμφέροντα.

Η ιδανική περίπτωση είναι το Grid να περιλαμβάνει αυξημένες δυνατότητες συνεργασίας, ή οποία όμως θα συνοδεύεται και από αντίστοιχο υψηλό επίπεδο ασφάλειας του συστήματος. Δυστυχώς όμως, μέχρι σήμερα, αποδεικνύεται ότι η επίτευξη μαζί των στόχων αυτών παρουσιάζει σημαντικές δυσκολίες. Αυτές αναμένεται να ξεπεραστούν, ως ένα βαθμό, στο μέλλον, με την εμφάνιση νέων βελτιωμένων πρωτοκόλλων ασφάλειας και συνεργασίας.

## **2.7 Σύγκριση των τεχνολογιών πλέγματος με άλλες τεχνολογίες**

Ο σχεδιασμός του Grid έχει βασιστεί σε ένα ανοιχτό σύνολο προτύπων και πρωτοκόλλων τα οποία επιτρέπουν την επικοινωνία μεταξύ ετερογενών και γεωγραφικά κατανεμημένων

συστημάτων, όπως για παράδειγμα το Open Grid Services Architecture (OGSA) (για το τελευταίο βλ. ενότητα 3.2).

Η δομή του Grid θεωρούμε ότι αποτελεί εξέλιξη άλλων δομών και τεχνολογιών, όπως τα κατακεκομημένα συστήματα (distributed systems), ο παγκόσμιος ιστός (Web), οι υπολογισμοί ομοτίμων πόρων (peer-to-peer computing) και οι τεχνολογίες δυναμικής εικονικής οργάνωσης (dynamic virtualization). Λόγω αυτού, η τεχνολογία των συστημάτων Grid παρουσιάζει πολλές ομοιότητες με άλλες τεχνολογίες (στις οποίες βασίστηκε) αλλά και πολλές αντιθέσεις. Αναλυτικότερα:

- Όπως και στο Web, τα Grids κρύβουν όλη την πολυπλοκότητα, και οι χρήστες απολαμβάνουν μια απλή και υλοποιημένη υπηρεσία. Σε αντίθεση όμως με το Web το οποίο απλά χρησιμεύει για τη διευκόλυνση της επικοινωνίας, τα συστήματα Grid επιτρέπουν πλήρη συνεργασία, με στόχο την επίτευξη κοινών επιχειρηματικών στόχων.
- Όπως και τα P2P συστήματα, τα συστήματα Grid επιτρέπουν στους χρήστες να μοιράζονται αρχεία. Σε αντίθεση όμως με τα P2P συστήματα, επιτρέπουν συνεργασίες πολλών με πολλούς (many-to-many) και δεν περιορίζονται στα αρχεία, αλλά χρησιμεύουν για συναλλαγές κάθε είδους πόρων.
- Όπως και τα κατακεκομημένα συστήματα (ή clusters), τα Grids ενοποιούν υπολογιστικούς πόρους. Σε αντίθεση με τα κατακεκομημένα συστήματα (ή clusters) τα οποία χρειάζονται γεωγραφική εγγύτητα και ομοιογένεια των συστημάτων, τα συστήματα Grid μπορεί να είναι γεωγραφικά κατακεκομημένα και ετερογενή.
- Όπως και οι τεχνολογίες εικονικής οργάνωσης συστημάτων (virtualization), τα Grids βασίζονται στην εικονική οργάνωση (E.O.) πόρων. Σε αντίθεση με τις τεχνολογίες εικονικής οργάνωσης συστημάτων (virtualization) οι οποίες έχουν στόχο την οργάνωση και λειτουργία μια αυτόνομης εικονικής μηχανής, τα

συστήματα Grid επιτρέπουν την εικονική οργάνωση απέραντων, ανόμοιων, απομακρυσμένων πόρων.

## **2.8 Τα πλεονεκτήματα του πλέγματος**

Πολλές εταιρίες επιδιώκουν να αξιοποιήσουν τα πλεονεκτήματα της σημερινής δομής του Grid όσον αφορά τα οικονομικά και τα αποδοτικά οφέλη, χωρίς να περιορίζονται σε ένα σύστημα, που δεν θα αυξάνεται σύμφωνα με τις ανάγκες τους.

Για να παρέχουν στους πελάτες την λύση που χρειάζονται, οι επιχειρήσεις έχουν να λύσουν κυρίως προβλήματα που αφορούν την ασφάλεια και την ποιότητα των υπηρεσιών που παρέχονται στο πελάτη. Χρησιμοποιώντας τα πρότυπα που έχουν αναπτυχθεί, οι πελάτες τους έχουν τη δυνατότητα να κινηθούν σύμφωνα με την συμβατότητα με τις υπάρχουσες νέες τεχνολογίες Grid, αλλά και να υιοθετήσουν καινούργιες καθώς αυτές εξελίσσονται, προσφέροντας νέα πλεονεκτήματα και ευκολίες.

Η πλατφόρμα του Grid μπορεί να συνδυάζει τους περισσευούμενους πόρους σε ένα εταιρικό δίκτυο δημιουργώντας ένα ισχυρό Grid που μπορεί να μοιραστεί και να χρησιμοποιηθεί από ομάδες στην ίδια την εταιρία, ή ακόμη και σε γεωγραφικά απομακρυσμένες σε σχέση με την εταιρία ομάδες. Το πιο κοινό τεχνολογικό στοιχείο που θα μπορούσε να χρησιμοποιηθεί για αυτούς τους πόρους είναι οι ηλεκτρονικοί υπολογιστές γραφείου (desktop PCs), οι οποίοι συνήθως λειτουργούν χαμηλότερα από τις δυνατότητές τους. Συνήθως το ποσοστό χρήσης της υπολογιστικής τους ισχύς είναι στο 10% των δυνατοτήτων τους ακόμα και στις πρωτεύουσες επαγγελματικές τους εφαρμογές. Οργανώνοντας την υπολογιστική υποδομή σε επίπεδα, μπορεί να χρησιμοποιηθεί σαν έσοδο διαθέτοντας την για απαιτητικές εφαρμογές. Αυτό αποτελεί άμεσο όφελος για εταιρίες, που επιθυμούν να υλοποιήσουν τεχνολογίες Grid, χωρίς να περιορίζονται από μελλοντικές εξελίξεις και προοπτικές.

Συνοψίζοντας τα κύρια πλεονεκτήματα της τεχνολογίας Grid, κυρίως όσον αφορά την επιχειρηματική εφαρμογή της είναι:

- Χαμηλότερο κόστος υπολογισμού, καθώς η σχέση απόδοσης / τιμή είναι διαφορετική. Οι διάφορες πλατφόρμες Grid, ανάλογα με τον προσανατολισμό τους, δίνουν την δυνατότητα εκτέλεσης περισσότερων εργασιών με μικρότερες επενδύσεις σε διοίκηση και σε υλικό σε σχέση με τις συνηθισμένες που βασίζονται αποκλειστικά σε υλικό εξοπλισμό. Το Grid καταφέρνει να βελτιώσει τον λόγο απόδοση προς τιμή σε έναν οργανισμό εξαρτώμενα πάντα από τις διαστάσεις του οργανισμού αυτού.
- Γρηγορότερα ερευνητικά αποτελέσματα. Το πλεόνασμα ισχύος που προκύπτει από την πλατφόρμα του Grid, μπορεί να δώσει σε έναν οργανισμό την δυνατότητα για γρηγορότερα αποτελέσματα. Αυτό μάλιστα, αν πρόκειται για εταιρία, μπορεί να της δώσει τη δυνατότητα για να κάνει περισσότερες δουλειές και να κατακτήσει μεγαλύτερο μερίδιο αγοράς.
- Καλύτερα αποτελέσματα. Η αύξηση της ισχύος οδηγεί πολλές φορές σε αναθεώρηση των ερευνητικών προσανατολισμών και να δοκιμαστούν πολλά υποσχόμενες λύσεις, που προηγούμενα είχαν απορριφθεί λόγω περιορισμών σε χρόνο και χρήμα. Επίσης η ισχύς αυτή μπορεί να βοηθήσει ώστε να προκύψουν αποτελέσματα καλύτερης ποιότητας επιτρέποντας μεγαλύτερες αναλύσεις και δοκιμές.

## **2.9 Τομείς που ευνοούνται από την τεχνολογία πλέγματος**

Η τεχνολογία Grid έχει ήδη εφαρμοστεί σε ερευνητικό επίπεδο σε διάφορους τομείς και αναμένεται να υιοθετηθεί για την επίτευξη των στόχων που οριοθετούνται σε κάθε τομέα.

Αναλυτικότερα:

- Στον τομέα της ιατρικής, της βιολογίας και της γενετικής η τεχνολογία Grid μπορεί να παρέχει σημαντικές υπηρεσίες στη διάγνωση νοσημάτων και στην αξιόπιστη διάγνωση μέσω της επεξεργασίας ιατρικών εικόνων. Εφαρμογές, όπως η διάγνωση με τη βοήθεια ιατρικών απεικονίσεων και η ανάλυση πρωτεϊνών και βιομοριακών μεθόδων, βασίζονται σε μεγάλο βαθμό στην αυτοματοποιημένη συγκέντρωση και επεξεργασία μεγάλου όγκου δεδομένων σε πραγματικό χρόνο, καθώς και στην διανομή των αποτελεσμάτων της επεξεργασίας σε γεωγραφικά διασπαρμένες τοποθεσίες. Το πρόβλημα απαίτησης σε υπολογιστικούς πόρους γίνεται ακόμα εντονότερο αν ληφθεί υπόψη ότι για να επιτευχθεί η επεξεργασία με τη βοήθεια του μεγάλου όγκου των παραγόμενων δεδομένων, σε συνθήκες πραγματικού χρόνου, αυτά θα πρέπει να συσχετιστούν με επίσης μεγάλο όγκο ιστορικών δεδομένων. Η υιοθέτηση της τεχνολογίας Grid μπορεί να δώσει στις εφαρμογές αυτές τους πόρους που χρειάζονται έτσι ώστε να παράγουν χρήσιμα αποτελέσματα, γρήγορα και αποδοτικά.
- Οι αγορές χρήματος είναι από τους νέους τομείς που έχουν αρχίσει να υιοθετούν την τεχνολογία Grid. Τράπεζες, ασφαλιστικές εταιρείες, χρηματιστηριακές επιχειρήσεις και αναλυτές, όλοι έχουν τη δυνατότητα να βελτιώσουν την ποιότητα των υπηρεσιών που παρέχουν κάνοντας χρήση της τεχνολογίας Grid. Οι τεχνολογίες Grid μπορούν να παρέχουν μεγάλες υπηρεσίες στη μελέτη χρηματοοικονομικών αναλύσεων και επίλυσης σχετικών προβλημάτων όπως ταυτοποίηση ρίσκου σε επενδυτικά σχέδια, διαδικασίες, απαιτήσεις και κόστη, κατηγοριοποίηση, ποσοτικοποίηση, διαχείριση και ανίχνευση ρίσκου.
- Ο επιστημονικός και ο εκπαιδευτικός κλάδος αποτελούν αυτή τη στιγμή τους κύριους χρήστες της Grid τεχνολογίας. Εδώ υπάρχουν δύο κατηγορίες χρηστών του Grid. Η πρώτη περιλαμβάνει ερευνητικά ινστιτούτα και τμήματα που

διεξάγουν έρευνα πάνω στην επιστήμη των υπολογιστών. Η κατηγορία αυτή δεν υιοθετεί τις περισσότερες φορές τις εμπορικές λύσεις αλλά δημιουργεί τις δικές της. Τα μέλη της θεωρούν τους εαυτούς τους κυρίως ως παρέχοντες παρά ως καταναλωτές της Grid τεχνολογίας. Η δεύτερη κατηγορία, που σχετίζεται κυρίως με τη δημόσια έρευνα, θεωρείται κυρίως ως καταναλωτής της Grid τεχνολογίας. Τυπικά παραδείγματα εδώ, η NASA με τα διαστημικά της προγράμματα και η φυσική υψηλών ενεργειών, αντικείμενο έρευνας μεγάλων ινστιτούτων όπως το CERN. Άλλοι τομείς της επιστήμης, που ανήκουν στους συχνούς χρήστες της Grid τεχνολογίας, αποτελούν η χημεία και η μηχανική που περιλαμβάνουν εφαρμογές υπολογιστικά απαιτητικές ή εφαρμογές που στηρίζονται στην εξόρυξη δεδομένων (data mining).

- Ο βιομηχανικός και κατασκευαστικός τομέας αποτελεί αποδεδειγμένα ένα χώρο, όπου η εφαρμογή της Grid τεχνολογίας μπορεί να αποδειχθεί εξαιρετικά ωφέλιμη. Στην αυτοκινητοβιομηχανία, ο σχεδιασμός αυτοκινήτων και φορτηγών απαιτεί γρήγορους και βελτιστοποιημένους πόρους για να επιταχυνθεί. Στο κατασκευαστικό τομέα γίνεται σε μεγάλο βαθμό η χρήση υπολογιστικών προσομοιώσεων. Τυπικές εφαρμογές που ευνοούνται στην περίπτωση αυτή είναι η δυναμική των ρευστών, ο σχεδιασμός με τη βοήθεια υπολογιστή, ανάλυση πεπερασμένων στοιχείων και προσομοιώσεις συγκρούσεων.

# 3

## ***Υπηρεσιοστρεφείς αρχιτεκτονικές και πλέγματα: Η σημερινή κατάσταση***

Στο κεφάλαιο αυτό γίνεται αναφορά σε στοιχεία Αρχιτεκτονικής Σχεδίασης των υπηρεσιοστρεφών υποδομών. Αρχικά αναλύονται οι παράμετροι σχεδίασης και το κεφάλαιο συνεχίζει περιγράφοντας την Αρχιτεκτονική Ανοιχτών Υπηρεσιών Πλέγματος (OGSA) και την Υπηρεσιοστρεφή Αρχιτεκτονική (SOA).

### ***3.1 Υπηρεσιοστρεφής αρχιτεκτονική***

Η προσαρμογή κάποιων υπηρεσιών σε μια εφαρμογή δεν είναι μια δύσκολη διαδικασία και επιπρόσθετα αναμένεται να αποδώσει στην εφαρμογή πρόσθετα λειτουργικά χαρακτηριστικά. Τα χαρακτηριστικά αυτά όμως δεν δημιουργούν μια υπηρεσιοστρεφή αρχιτεκτονική (*Service Oriented Architecture - SOA*), αφού η διαφορά τους είναι πιο μεγάλη.

Το πρότυπο υπηρεσιοστρεφούς αρχιτεκτονικής είναι ένα σχεδιαστικό μοντέλο με κύριο χαρακτηριστικό την ενσωμάτωση λογικής εφαρμογών μέσα σε υπηρεσίες που θα αλληλεπιδρούν μέσω επικοινωνιακών πρωτοκόλλων. Βάσει αυτού, η υιοθέτηση σε μια εφαρμογή, SOA δομής σημαίνει αυτόματα την αποδοχή κάποιων σχεδιαστικών αρχών και πρόσθετων τεχνολογιών ως βασικού τμήματος του τεχνικού περιβάλλοντος της.

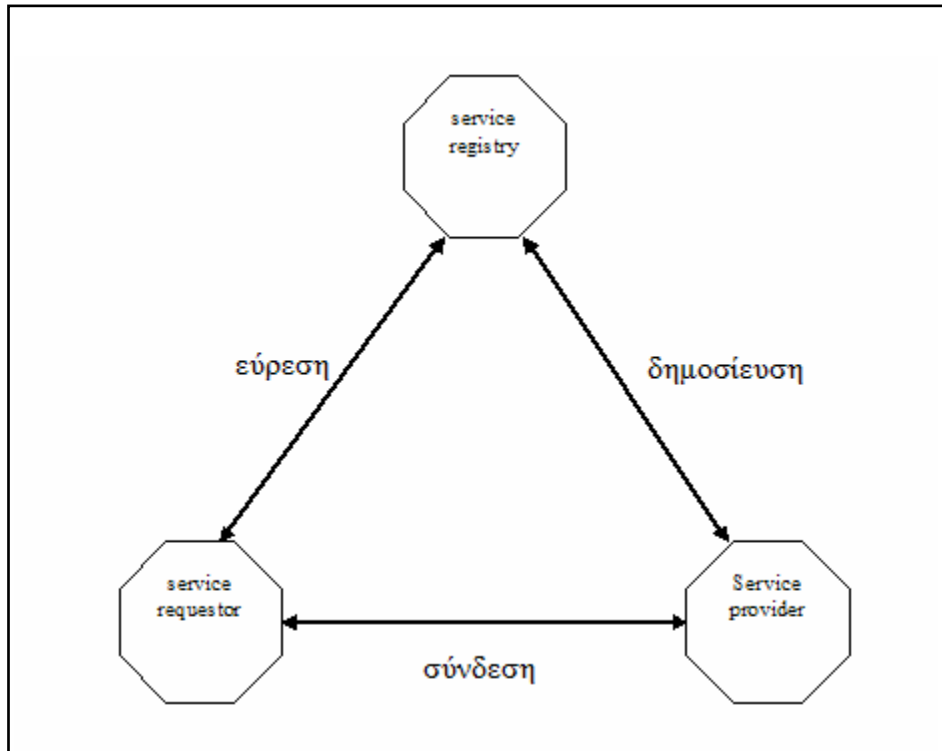
Σε επίπεδο σχεδιασμού συστημάτων, η χρήση της τεχνολογίας των Web Services οδηγεί στην υιοθέτηση της λεγόμενης Service-Oriented αρχιτεκτονικής. Οι βασικοί ρόλοι και λειτουργίες στην αρχιτεκτονική αυτή παρουσιάζονται στο Σχήμα 4. Η αρχιτεκτονική αυτή υποδεικνύει μια σχέση εξυπηρετητή-πελάτη (server-client) ανάμεσα στον προμηθευτή υπηρεσιών (service provider, που παίζει το ρόλο του server) και τον ζητώντα την υπηρεσία (service-requestor, που παίζει το ρόλο του client). Ο service-provider είναι αυτός που παρέχει την υπηρεσία δεχόμενος μηνύματα κλήσεις από τους requestors. Είναι επίσης υπεύθυνος για τη δημιουργία της περιγραφής της υπηρεσίας (service description) και τη δημοσίευσή της σε κάποιο κατάλογο-οδηγό υπηρεσιών (Universal Description, Discovery and Integration- UDDI) [15]. Ο service requestor αναζητά την υπηρεσία και την περιγραφή της σε κάποιο κατάλογο υπηρεσιών (service registry) και στη συνέχεια καλεί κατάλληλα την επιθυμητή υπηρεσία. Ο κατάλογος υπηρεσιών φέρνει ουσιαστικά τις δύο πλευρές, client και server σε επαφή. Η συνέχεια αφορά μόνο τις δύο άλλες συμμετέχουσες μονάδες (service requestor και service provider).

### **3.1.1 Γενικά για τις υπηρεσίες**

Το τελευταίο καιρό έχει γίνει αρκετά μεγάλη συζήτηση γύρω από το θέμα των υπηρεσιών των εφαρμογών. Οι υπηρεσίες τείνουν να γίνουν τμήματα της εφαρμογής που αθροιστικά σχηματίζουν το περιβάλλον αυτής. Φυσικά δεν αποτελούν απλά ένα κομμάτι της εφαρμογής, αλλά έχουν χαρακτηριστικά που τις μετατρέπουν σε μέρος μιας αρχιτεκτονικής προσανατολισμένης στις υπηρεσίες (Service Oriented Architecture).

Ένα από τα χαρακτηριστικά αυτά είναι η αυτονομία από άλλες υπηρεσίες. Αυτό σημαίνει ότι κάθε υπηρεσία είναι υπεύθυνη για το δικό της εύρος λειτουργίας, περιορίζοντας έτσι και εξειδικεύοντάς την σε συγκεκριμένες επαγγελματικές χρήσεις. Αυτός ο σχεδιασμός έχει ως αποτέλεσμα τη δημιουργία ανεξάρτητων μονάδων, ελαστικά συνδεδεμένων μεταξύ τους με κάποιο πρότυπο πλαίσιο επικοινωνίας. Εξαιτίας αυτής της ανεξαρτησίας





Σχήμα 4: Δομή υπηρεσιοστρεφούς αρχιτεκτονικής

που απολαμβάνουν οι υπηρεσίες στο πλαίσιο αυτό, η προγραμματιστική λογική που κάθε υπηρεσία χρησιμοποιεί, δεν χρειάζεται να προσαρμόζεται σε συγκεκριμένη πλατφόρμα ή τεχνολογία. Είναι χαρακτηριστικό των υπηρεσιών ενός πλαισίου, ο πολυμορφισμός των μερών του με ταυτόχρονη ομαλή συνεργασία.

### 3.1.2 Διαδικτυακές υπηρεσίες

Ο πιο ευρέως διαδεδομένος και επιτυχημένος τύπος υπηρεσιών είναι οι διαδικτυακές υπηρεσίες XML Services γνωστές ως Web Services. Αυτός ο τύπος υπηρεσίας έχει δύο βασικές προαπαιτήσεις:

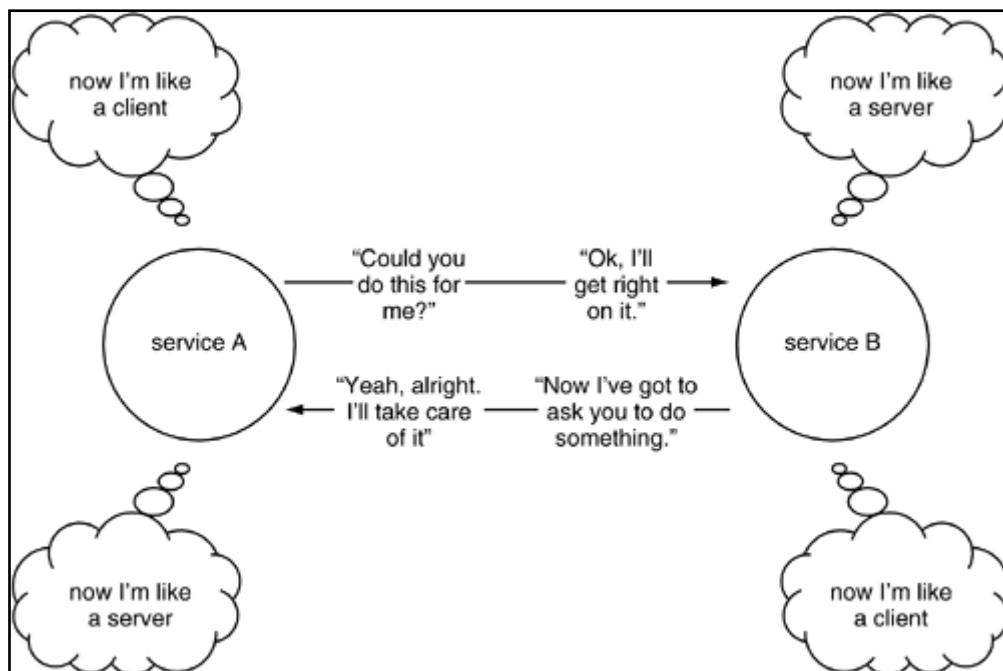
- επικοινωνεί μέσω πρωτοκόλλου Internet (κυρίως HTTP)
- στέλνει και δέχεται δεδομένα μέσα από XML αρχεία

*Η σελίδα αυτή είναι σκόπιμα λευκή*

Η ευρεία αποδοχή του μοντέλου των Web Services είχε ως αποτέλεσμα την ανάγκη πρόσθετων τεχνολογιών που βασίζονται σε αυτές και τη δημιουργία καινούργιων προτύπων. Έτσι η “παραγωγή” τέτοιων υπηρεσιών απαιτεί:

- τη περιγραφή της υπηρεσίας αναλυτικά, τουλάχιστον με ένα WSDL έγγραφο
- τη δυνατότητα μεταφοράς ενός XML εγγράφου χρησιμοποιώντας SOAP μέσω HTTP.

Επιπρόσθετα είναι συνηθισμένο μια υπηρεσία να λειτουργεί και ως πελάτης (client/requestor) και ως πάροχος (provider) υπηρεσίας. Αναλόγως λοιπόν με τη δραστηριότητα της υπηρεσίας κάθε στιγμή, μετατρέπεται από το ένα στο άλλο. Στο παρακάτω σχήμα φαίνεται ένα παράδειγμα αυτής της συμπεριφοράς:



Σχήμα 5: Ανταλλαγή ρόλων μεταξύ Web services κατά τη διάρκεια μιας επικοινωνίας

### 3.1.3 Web Services Description Language (WSDL)

Οι Web Services χρειάζεται να ορίζονται με συγκεκριμένο τρόπο έτσι ώστε να μπορούν να εντοπιστούν και χρησιμοποιηθούν από άλλες υπηρεσίες και εφαρμογές. Για αυτό το

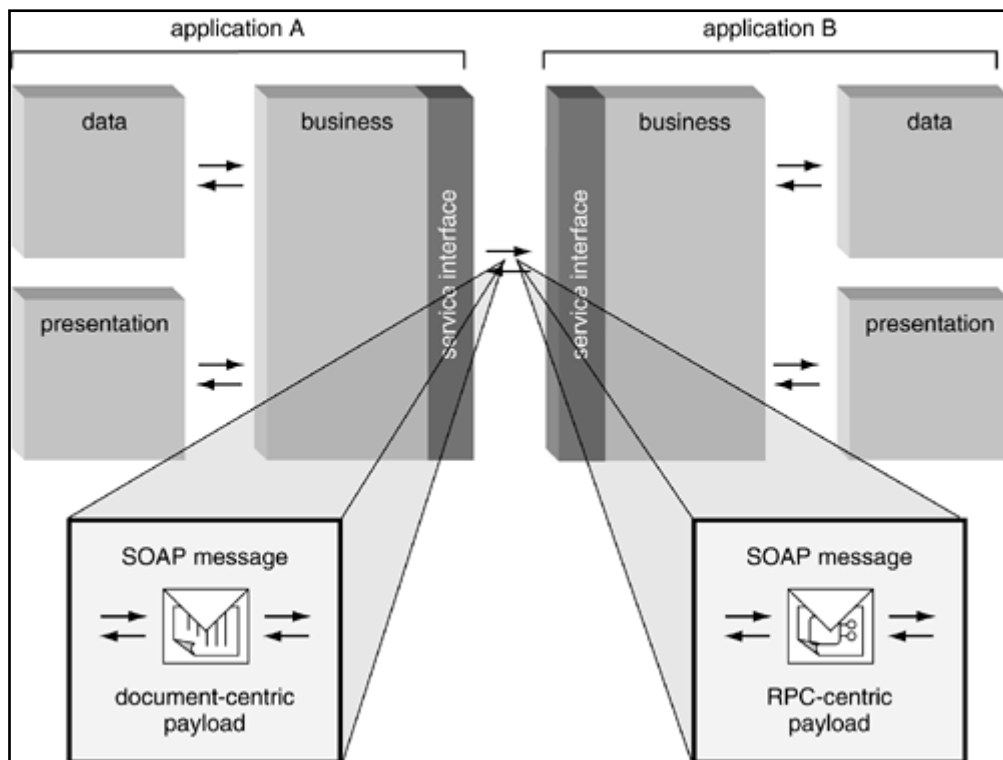
*Η σελίδα αυτή είναι σκόπιμα λευκή*

σκοπό δημιουργήθηκε από τον οργανισμό W3C, μια γλώσσα περιγραφής γνωστή ως *Web Services Description Language (WSDL)* [31].

Η γλώσσα αυτή είναι μια XML δομή εγγράφου, που περιέχει όλες τις πληροφορίες σχετικά με τα δεδομένα εισόδου και εξόδου, τις μεθόδους και ότι άλλο χρειάζεται να γνωρίζει κάποιος για την ακριβή χρήση μιας διαδικτυακής υπηρεσίας.

### 3.1.4 Simple Object Access Protocol (SOAP)

Αν και αρχικά είχε θεωρηθεί ως η τεχνολογία που θα γεφυρώσει το κενό μεταξύ ανόμοιων πλατφορμών βασισμένων σε RPC επικοινωνία, το SOAP [83] έχει εξελιχθεί στο ευρέως χρησιμοποιούμενο πρότυπο επικοινωνίας για τη χρήση XML Υπηρεσιών Διαδικτύου (Web Services). Μετά από αυτή την κατάσταση γίνεται πολλές φορές η παράφραση του ακρωνύμιου από Simple Object Access Protocol σε Service-Oriented Architecture (or Application) Protocol.



Σχήμα 6: Χρήση SOAP Πρωτοκόλλου σε εφαρμογές

*Η σελίδα αυτή είναι σκόπιμα λευκή*

Το πρωτόκολλο SOAP διαμορφώνει ένα πρότυπο μήνυμα που αποτελείται από ένα XML έγγραφο ικανό να περιγράψει δεδομένα όπως RPC κλήσεις κ.α. Το μήνυμα αυτό μεταφέρεται μεταξύ των υπηρεσιών και των εφαρμογών, χρησιμοποιώντας κυρίως το HTTP πρωτόκολλο δικτύου. Με τον τρόπο αυτό ολοκληρώνεται το πλαίσιο λειτουργίας και επικοινωνίας στην SOA δομή, αφού με την βοήθεια της περιγραφής WSDL είναι εφικτή η επικοινωνία και συνεργασία οποιωνδήποτε υπηρεσιών στο δίκτυο.

Στο σχήμα (Σχήμα 6) βλέπουμε τη χρήση του πρωτοκόλλου σε εφαρμογές είτε για προτυποποιημένη RPC επικοινωνία είτε για γενική χρήση μεταφοράς μηνύματος.

Το πρωτόκολλο SOAP καθορίζει ένα XML έγγραφο με μια συγκεκριμένη δομή που μέσα του θα εμπεριέχονται οι πληροφορίες. Η βασική δομή ενός τέτοιου εγγράφου είναι η παρακάτω:

```
<soap:Envelope xmlns:soap="http://schemas.xmlsoap.org/soap/envelope/">
  <soap:Body>
    ...
  </soap:Body>
</soap:Envelope>
```

**Σχήμα 7: Βασική δομή μηνύματος SOAP**

Εσωτερικά του “σώματος” (Body) του εγγράφου, ενσωματώνεται η πληροφορία που θέλουμε να μεταφέρουμε από τη μία υπηρεσία στην άλλη. Η τεχνολογία που χρησιμοποιείται για απομακρυσμένη κλήση υπηρεσιών είναι η SOAP-RPC, (επέκταση της τεχνολογίας Remote Procedure Call) η οποία εισάγει στο SOAP έγγραφο τις απαιτούμενες πληροφορίες, σύμφωνα και με το WSDL της υπηρεσίας που θα κληθεί και στέλνει το μήνυμα. Επιπλέον, όπως φαίνεται και στο Σχήμα 6, εκτός από την χρήση που περιγράφηκε, το SOAP μπορεί να μεταφέρει οποιαδήποτε άλλη πληροφορία ενσωματωμένη σε XML δομή, μέσα στο “φάκελο” του SOAP μηνύματος, αρκεί βέβαια ο παραλήπτης να γνωρίζει πώς να αποκωδικοποιήσει το έγγραφο αυτό.

### **3.2 Αρχιτεκτονική ανοιχτών υπηρεσιών πλέγματος**

Οι συντονισμένες μελέτες του *Open Grid Forum* (OGF) - ενός οργανισμού αποτελούμενου από χρήστες, προγραμματιστές, ερευνητές με σκοπό την ανάπτυξη προτύπων για την υποστήριξη των Grid τεχνολογιών - κατέληξε το 2002 στην Αρχιτεκτονική Ανοιχτών Υπηρεσιών - *Open Grid Service Architecture* [73], η οποία συντάχθηκε σε ένα κείμενο που περιγράφει την δομή ενός Grid συστήματος και συγκεντρώνει όλα τα πρότυπα που χρησιμοποιεί η αρχιτεκτονική αυτή σε κάθε επίπεδο της.

#### **3.2.1 Στόχοι της αρχιτεκτονικής**

Οι κύριοι στόχοι του OGSA είναι οι ακόλουθοι:

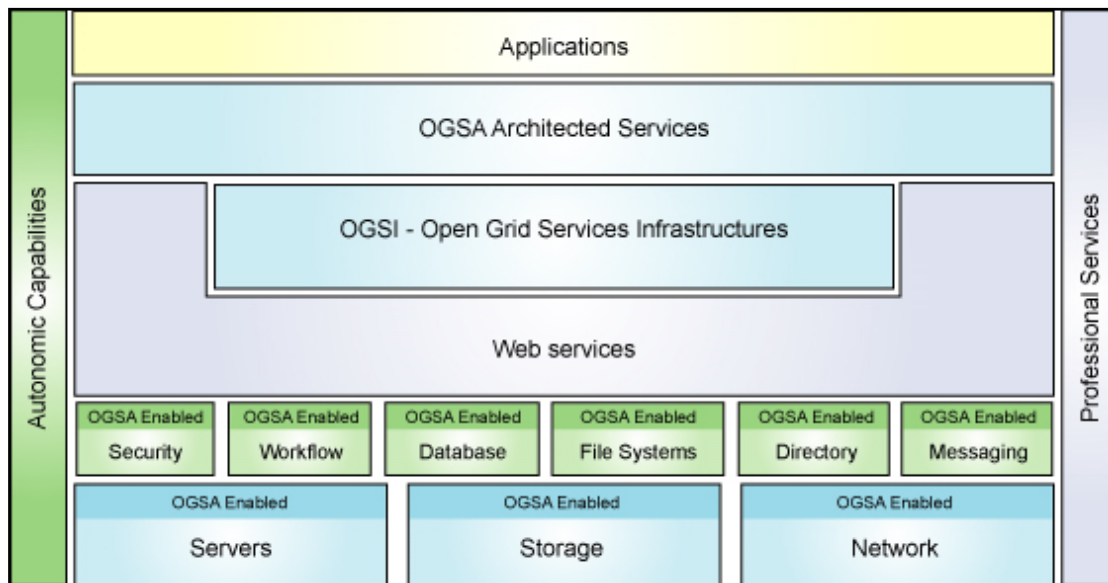
- να διαχειρίζεται πόρους απομακρυσμένων ετερογενών συστημάτων
- να αποδίδει σημαντική ποιότητα στις υπηρεσίες που παρέχει (Quality of Service)
- να παρέχει τις βάσεις για αυτόνομες λύσεις διαχείρισης. Η διαφορετικότητα των πόρων που συμμετέχουν σε ένα Grid και ο δυναμικός τρόπος συνεργασίας τους απαιτεί ένα σύστημα διαχείρισης που θα προσαρμόζεται σύμφωνα με τις εκάστοτε ανάγκες.
- να καθορίζει γνωστά πρότυπα και πρωτόκολλα λειτουργίας. Η δυναμική εισαγωγή και συνεργασία ετερογενών συστημάτων στο δίκτυο του Grid, βασίζεται στην υιοθέτηση γνωστών, ευρείας χρήσης προτύπων.
- να εκμεταλλεύεται υπάρχουσες τεχνολογίες και να τις προσαρμόζει στο Grid, αν είναι εφικτό. Το OGSA βασίζεται στη τεχνολογία των Web Services (υπηρεσίες διαδικτύου) όπως αυτές αναφέρονται στην ενότητα 3.1.2.



### 3.2.2 Περιγραφή της αρχιτεκτονικής

Το OGSA αποτελείται από τέσσερα βασικά επίπεδα (όπως φαίνεται και στο σχήμα που ακολουθεί – Σχήμα 8):

- τους πόρους (φυσικούς και λογικούς)
- τις υπηρεσίες διαδικτύου (Web Services) συμπεριλαμβανομένου και της υποδομής Open Grid Services Infrastructure (OGSI) που έχει άμεση σχέση με τις υπηρεσίες
- τις υπηρεσίες σχεδιασμένες από το OGSA
- τις Grid εφαρμογές



Σχήμα 8: Αρχιτεκτονική ανοιχτών υπηρεσιών πλέγματος

#### 3.2.2.1 Επίπεδο φυσικών και λογικών πόρων

Αυτό είναι το πιο χαμηλό επίπεδο από τα τέσσερα και έχει να κάνει με τους πόρους που συμμετέχουν στο Grid. Η έννοια του 'πόρου' στο OGSA είναι πολύ σημαντική και δεν είναι συγκεκριμένη. Ως πόρος (resource) σε ένα Grid μπορεί να θεωρηθεί από τον επεξεργαστή ενός υπολογιστή, μέχρι το τμήμα υπολογιστών μιας μεγάλης εταιρίας. Επίσης εκτός από υπολογιστικές μονάδες, συμμετέχουν και μονάδες αποθήκευσης, βάσεις

*Η σελίδα αυτή είναι σκόπιμα λευκή*

δεδομένων, δικτυακές υποδομές κλπ. Αυτά που αναφέρθηκαν παραπάνω είναι κυρίως οι φυσικοί πόροι. Εκτός από αυτούς έχουμε και τους λογικούς πόρους, κάποια ενδιάμεσα συστήματα (middleware) που χρησιμοποιώντας τους φυσικούς πόρους, προσφέρουν στοιχειώδεις υπηρεσίες, όπως διαχείριση αρχείων ή βάσεων δεδομένων, στο παραπάνω επίπεδο.

#### **3.2.2.2 Επίπεδο υπηρεσιών διαδικτύου (Web Services)**

Μία πολύ βασική θεώρηση του OGSA είναι ότι όλοι οι πόροι του συστήματος αντιστοιχίζονται με υπηρεσίες. Έτσι η OGSΙ υποδομή (Open Grid Services Infrastructure) χρησιμοποιεί συγκεκριμένες, γνωστές υπηρεσίες δικτύου και πρωτόκολλα όπως XML και WSDL για να δημιουργήσει επαφές και διασυνδέσεις κάθε Grid υπηρεσίας με τους πόρους του συστήματος. Η OGSΙ επεκτείνει τις δυνατότητες των δικτυακών υπηρεσιών έτσι ώστε να παρέχει δυναμικές και αξιόπιστες υπηρεσίες όπως απαιτείται για τη μοντελοποίηση των πόρων.

#### **3.2.2.3 Επίπεδο Grid υπηρεσιών**

Οι υπηρεσίες διαδικτύου και οι δυνατότητες του OGSΙ παρέχουν τη βασική υποδομή για το επόμενο επίπεδο, των Grid υπηρεσιών. Αυτή την εποχή υπάρχει μεγάλη δραστηριότητα προς την κατεύθυνση προσδιορισμού και προτυποποίησης τέτοιων υπηρεσιών, όπως υπηρεσίες υπολογισμού, πληροφοριών, πυρήνα, κ.α. Καθώς οι υλοποιήσεις αυτών των υπηρεσιών θα αρχίσουν να εμφανίζονται, η OGSA αρχιτεκτονική θα γίνεται όλο και πιο χρήσιμη βασισμένη πάντα στις υπηρεσίες, πρόκειται δηλαδή για Service Oriented Architecture, όπως περιγράφεται στην ενότητα 3.1.

#### **3.2.2.4 Επίπεδο εφαρμογών Grid**

Με το πέρασμα του χρόνου και όσο οι Grid υπηρεσίες θα αναπτύσσονται, καινούργιες εφαρμογές βασισμένες στην Grid τεχνολογία θα κάνουν την εμφάνισή τους. Αυτές οι

εφαρμογές θα χρησιμοποιούν μια ή και περισσότερες Grid υπηρεσίες του προηγούμενου επιπέδου.

Η χρήση υπηρεσιοστρεφών αρχιτεκτονικών σχεδιασμός αποτελεί την επικρατούσα τάση για τα νέα υπολογιστικά συστήματα και σε αυτό το πλαίσιο ένας μεγάλος αριθμός από WS-\* πρότυπα έχει αναπτυχθεί σα σημείο αναφοράς για τα νέα αυτά συστήματα.

# 4

## ***Ποιότητα υπηρεσιών και επιχειρηματικά μοντέλα***

### ***4.1 Ποιότητα υπηρεσιών: Η σημερινή κατάσταση***

Η ετερογένεια και η δυναμική του πλέγματος και των υπηρεσιοστρεφών υποδομών αναφορικά με την πολλαπλότητα των υπηρεσιών, των χρηστών, και των απαιτήσεών τους επιτρέπει σε όλες τις εμπλεκόμενες οντότητες (πόρους, χρήστες, παρόχους, ενδιάμεσους μεταπωλητές) να λειτουργούν αυτόνομα. Τα ιδιαίτερα τεχνολογικά και επιχειρηματικά χαρακτηριστικά του πλέγματος και των υπηρεσιοστρεφών υποδομών αναδεικνύουν την διασφάλιση της ποιότητας υπηρεσιών επιτακτική ανάγκη [53].

#### ***4.1.1 Περιγραφή αίτησης για ποιότητα υπηρεσιών***

Συνήθως το αίτημα για την υπηρεσία υποβάλλεται από κοινού με τις αντίστοιχες ανάγκες για ποιότητα υπηρεσίας καθώς η ποιότητα της υπηρεσίας είναι αυτό που χαρακτηρίζει τον τρόπο και τις συνθήκες που αυτή θα προσφέρεται και κατά συνέπεια και την ίδια την υπηρεσία. Σαν αποτέλεσμα ο μηχανισμός υποβολής αιτήσεων υπηρεσίας και ποιότητας υπηρεσιών πρέπει να είναι όχι μόνο ιδιαίτερα εύκαμπτος και εκφραστικός για να ικανοποιήσει τις προτιμήσεις των διάφορων χρηστών αλλά επίσης απλός και εύχρηστος ώστε να μπορούν οι χρήστες εκφράσουν με αποδοτικό τρόπο τις ανάγκες τους.

### **4.1.2 Περιγραφή υπηρεσιών**

Το πλέγμα και οι υπηρεσιοστρεφείς υποδομές είναι εξαιρετικά ανοικτά και δυναμικά περιβάλλοντα με τις υπηρεσίες τους συχνά προστιθέμενες και προσφερόμενες σε διαφορές παραλλαγές. Για τις υπηρεσίες ιστού (web services), οι μηχανισμοί περιγραφής WSDL είναι λύση σε αυτό πρόβλημα. Δυστυχώς όμως οι μηχανισμοί αυτοί εστιάζουν μόνο στο πώς κάποιος μπορεί να έχει πρόσβαση στην υπηρεσία και τι ακριβώς προσφέρει η υπηρεσία, χωρίς να περιλαμβάνει άλλες πληροφορίες που είναι χρήσιμες για τον καθορισμό της ποιότητας των υπηρεσιών. Κατά συνέπεια ο μηχανισμός περιγραφής WSDL είναι ανεπαρκής και ένας νέος κάποιος θα υποβληθεί. Η λύση σε αυτό είναι συμπληρωματικοί μηχανισμοί που θα καθορίζουν επακριβώς το επίπεδο παροχής ποιότητας υπηρεσιών σε επιχειρηματικό επίπεδο που θα εκφράζει όχι μόνο τέτοιο παραδοσιακό QoS παράγοντες ως καθαρό πλάτος, αλλά και τι δυνατότητες χρήσης, ασφάλεια, και άλλους παράγοντες που είτε ζητούνται από το χρήστη είτε προσφέρονται από την υπηρεσία.

### **4.1.3 Προσαρμοστική ποιότητα υπηρεσιών**

Στο μοντέλο λειτουργίας του πλέγματος και των υπηρεσιοστρεφών υποδομών, οι υπηρεσίες μπορούν να είναι αλλάζουν ή να διαγράφονται δυναμικά και οποιαδήποτε στιγμή, ενώ νέες υπηρεσίες ή συνδυασμός τους μπαίνουν στο «οικοσύστημα». Επιπλέον χαρακτηριστικά όπως η τοπολογία δικτύων και οι πλατφόρμες λειτουργίας μπορούν αυτόνομα να αλλάζουν και να προσαρμοστούν στις ανάγκες των χρηστών. Σε αυτό το «ρευστό» περιβάλλον απαιτούνται μηχανισμοί προκειμένου να εντοπιστούν εγκαίρως οι αλλαγές και να το σύστημα να προσαρμοστεί κατάλληλα ώστε να διασφαλιστεί στον τελικό χρήστη η ποιότητα υπηρεσιών.

#### **4.1.4 Δέσμευση πόρων και δρομολόγηση**

Συνήθως στην αίτηση ενός χρήστη για ποιότητα υπηρεσίας μπορούν να ανταποκριθούν περισσότεροι του ενός πόροι με αποτέλεσμα το πλέγμα ή η υπηρεσιοστρεφής υποδομή να πρέπει κάθε φορά να δεσμεύσει τους πιο κατάλληλους για να διασφαλίσουν όχι μόνο την ποιότητα υπηρεσιών για τον συγκεκριμένο χρήστη αλλά και για όλο το πλέγμα ή την υποδομή γενικά. Σε αυτό το πλαίσιο, η εξισορρόπηση φόρτου εργασίας και η διαπραγμάτευση των πόρων είναι απαραίτητη [99]. Από τη μία πλευρά ο σχεδιασμός της πολιτικής που στοχεύει να φθάσει στην ποιότητα υπηρεσίας που μπορεί να επιτευχθεί περιλαμβάνει κάποιους συμβιβασμούς και από την άλλη ή συνολική δέσμευση και δρομολόγηση πόρων πρέπει να εξασφαλίζει την πλήρη αξιοποίηση των πόρων του πλέγματος ή της υποδομής.

#### **4.1.5 Αξιολόγηση παροχής ποιότητας υπηρεσιών**

Οι επιδόσεις των υπηρεσιών σε αυτά τα περιβάλλοντα πρέπει να αξιολογούνται συνεχώς με σκοπό την μεγαλύτερη ικανοποίηση των χρηστών και τη βελτίωση από πλευράς παρόχου των μηχανισμών παροχής ποιότητας υπηρεσιών. Με αυτό τον τρόπο τόσο οι χρήστες μπορούν να επιλέξουν πιο αποδοτικά τους παρόχους υπηρεσιών εξασφαλίζοντας μακροπρόθεσμα υψηλότερη ποιότητα υπηρεσιών ενώ οι πάροχοι μπορούν να βελτιώνουν τις υπηρεσίες τους και τη χρήση των πόρων τους.

## **4.2 Επιχειρηματικά μοντέλα**

Η παροχή υπηρεσιών σχετίζεται άμεσα με τον τρόπο λειτουργίας της υποδομής μέσα ή μέσω της οποίας προσφέρεται και σχετίζεται με τα επιχειρηματικά μοντέλα των υποδομών αλλά και με τις τεχνολογίες με τις οποίες αυτά υλοποιούνται. Στην περίπτωση των ακαδημαϊκών / ερευνητικών πλεγμάτων η ποιότητα υπηρεσίας είναι η καλύτερη που μπορεί να προσφέρει την παρούσα χρονική στιγμή (best effort) [3] ενώ δεν υπάρχουν ούτε συμφωνίες για το επίπεδό της ούτε πρόστιμα αν αυτή δεν ικανοποιεί τον τελικό χρήστη.

Από την άλλη στις υποδομές με επιχειρηματικό προσανατολισμό, όπως το Amazon Elastic Computing Cloud [7], η ποιότητα υπηρεσιών καθορίζεται από τις συμφωνίες μεταξύ χρήστη και παρόχου (Service Level Agreement – SLA) οι οποίες περιλαμβάνουν μεταξύ άλλων τους τρόπους διαχείρισης της ποιότητας υπηρεσιών και τις πιθανές αποζημιώσεις αν υπηρεσίες δεν προσφέρονται με τον τρόπο που έχει προσυμφωνηθεί.

#### **4.2.1 Τι είναι τα επιχειρηματικά μοντέλα;**

Οποιαδήποτε συζήτηση για επιχειρηματικά μοντέλα (Business Models) [63] περιπλέκεται αμέσως από την έλλειψη κοινού ορισμού και κατανόησης της έννοιας. Μια αναζήτηση του όρου στο Google επιστρέφει αρκετά εκατομμύρια αναφορές, πολύ λίγες από τις οποίες μοιράζονται έναν κοινό ορισμό ή δομή για την έννοια. Ως εκ τούτου υπάρχει η ανάγκη να καθοριστεί μια κοινή κατανόηση του όρου, και αυτό πρέπει να βασιστεί στην καλά-αποδεκτή φρόνηση.

Η συνηθέστερα χρησιμοποιημένη περιγραφή του όρου είναι απλά ότι το επιχειρηματικό μοντέλο αναλύει πώς η επιχείρηση «κάνει χρήματα». Η μικρότερη διαφορά σε ένα μικρό στοιχείο ανάμεσα σε δύο επιχειρηματικό μοντέλο μπορεί να σημάνει την επιτυχία του ενός και την αποτυχία άλλου. Μεταξύ των συχνότερα αναφορών είναι εκείνοι από τον Timmers [76], ο οποίος περιγράφει ένα επιχειρηματικό μοντέλο όπως:

- Μια αρχιτεκτονική για τις ροές προϊόντων, υπηρεσιών και πληροφοριών, συμπεριλαμβανομένης μιας περιγραφής των διάφορων επιχειρησιακών δραστηριοτήτων και των ρόλων τους,
- Μια περιγραφή των πιθανών οφελών για τους διάφορους επιχειρησιακούς δράστες και
- Μια περιγραφή των εισοδηματικών πηγών.



Οι Chesborough και Rosenbloom [46] διευκρινίζουν ότι οι έξι λειτουργίες ενός επιχειρησιακού προτύπου είναι:

- Ανάλυση της πρότασης αξίας, δηλ. η αξία που δημιουργείται για τους χρήστες με την προσφορά.
- Προσδιορισμός του τομέα αγοράς, δηλ. οι χρήστες στους οποίους η προσφορά είναι χρήσιμη και για ποιο σκοπό, όπως και ο προσδιορισμός του μηχανισμού παραγωγής εισοδήματος για την εταιρία.
- Καθορισμός α) της δομής της αλυσίδας προστιθέμενης αξίας (value chain) για να κατανεμηθεί η προσφορά και β) των συμπληρωματικών προτερημάτων που απαιτούνται για να υποστηρίξουν τα προτερήματα της εταιρίας σε αυτήν την αλυσίδα.
- Ανάλυση του κόστους και των δυνατοτήτων κέρδους για την παραγωγή του προϊόντος δεδομένης της δομής αλυσίδων προστιθέμενης αξίας που έχει επιλεχθεί.
- Περιγραφή της θέσης της εταιρίας μέσα στους προμηθευτές και τους πελάτες των δικτύων αξίας (value networks) συμπεριλαμβανομένου του προσδιορισμού των πιθανών ανταγωνιστών.
- Διατύπωση της ανταγωνιστικής στρατηγικής από την οποία η καινοτόμος εταιρία θα αποκομίσει και θα κρατήσει το πλεονέκτημα πέρα από τους ανταγωνιστές.

Δεν υπάρχει κανένα συμφωνηθέν σχήμα για την παρουσίαση των επιχειρησιακών προτύπων ενώ το πιο αποδεκτό σχήμα είναι μια ενιαία σελίδα σε απλή αφηγηματική που εξετάζει τα ανωτέρω θέματα. Πρέπει να σημειωθεί ότι το επιχειρηματικό μοντέλο δεν περιλαμβάνει ρητά τις τεχνολογίες εφαρμογής, οδηγείται από τις επιχειρησιακές απαιτήσεις ενώ απλά η τεχνολογία επιτρέπει την υλοποίησή του. Επίσης η άρθρωση ενός

επιχειρησιακού προτύπου δεν αποτελεί σε καμία περίπτωση εγγύηση της επιτυχίας της και μόνο η ίδια εφαρμογή θα καθορίσει εάν το πρότυπο λειτουργήσει ή όχι. Τέλος να σημειωθεί ότι τα νέα επιχειρηματικά μοντέλα αντιμετωπίζονται όλο και περισσότερο ως βασική πνευματική ιδιοκτησία, και συχνά οι προσπάθειες γίνονται να τα προστατεύσουν από το δίπλωμα ευρεσιτεχνίας, ενώ πολύ λιγότερο αντιμετωπίζονται ως εμπορικά μυστικά.

Με βάση τα παραπάνω λοιπόν, για το καθορισμό ενός επιχειρηματικού μοντέλου και έπειτα την ποιότητα υπηρεσίας καλούμαστε να απαντήσουμε στα παρακάτω ερωτήματα:

- Ποια είναι η προσφορά της υπηρεσίας στον πελάτη;
- Ποια αξία προστίθεται με την ενσωμάτωση της εφαρμογής σε ένα περιβάλλον πλέγματος;
- Ποιο είναι το κόστος για την προσφορά της υπηρεσίας στον πελάτη; Πώς καθορίζονται οι τιμές;
- Πώς οι πελάτες θα πληρώσουν για τις υπηρεσίες;
- Πώς παρέχεται μία υπηρεσία;

Έχοντας ορίσει το πλαίσιο του επιχειρηματικού μοντέλου, θα προκύψουν οι λειτουργικές απαιτήσεις της υποδομής που θα το κάνει πραγματικότητα και οι μηχανισμοί που θα είναι σε θέση να διασφαλίσουν την απαιτούμενη ποιότητα υπηρεσιών.

#### **4.2.2 Οι ρόλοι σε ένα επιχειρηματικό μοντέλο**

Κάθε επιχειρηματικό μοντέλο χρειάζεται τους δράστες, και αυτοί οι δράστες διαδραματίζουν τους ρόλους οι οποίοι καθορίζουν την ευρεία ευθύνη των δραστών. Εδώ περιγράφουμε τους διαφορετικούς ρόλους που ισχύουν στα επιχειρηματικά μοντέλα πλέγματος και υπηρεσιοστρεφών υποδομών. Πρέπει να σημειώσουμε ότι μια

οργάνωση/δράστης μπορεί να διαδραματίσει περισσότερους από έναν ρόλους συγχρόνως, και στα ίδια και σε διαφορετικά επιχειρηματικά μοντέλα. Μερικοί ρόλοι είναι γενικοί, άλλοι είναι πιο συγκεκριμένοι για την υποδομή ενώ σε πολλές περιπτώσεις οι ρόλοι συνδέονται έντονα με την παροχή υπηρεσίας και το είδος της. Για αυτό εδώ θεωρούμε τον όρο «υπηρεσία» με την ευρύτερη δυνατή έννοια.

#### **4.2.2.1 Προμηθευτής**

Αυτό είναι ένας γενικός όρος για οποιοδήποτε είδος του δράστη που παρέχει ένα αγαθό ή μια υπηρεσία. Συχνά αυτό είναι σε αντάλλαγμα για τα χρήματα.

#### **4.2.2.2 Πελάτης**

Ο πελάτης έχει επιχειρησιακή σχέση με έναν προμηθευτή που καθορίζεται από την συμφωνία επιπέδου υπηρεσιών (Service Level Agreement – SLA).

#### **4.2.2.3 Καταναλωτής**

Ο καταναλωτής είναι ο τελικός χρήστης της υπηρεσίας. Ο καταναλωτής μπορεί (αλλά όχι απαραίτητως) να είναι ο ίδιος δράστης με έναν πελάτη.

#### **4.2.2.4 Προμηθευτής πόρων**

Ο προμηθευτής πόρων είναι η οργάνωση που παρέχει το υλικό των πόρων (hardware), παραδείγματος χάριν υπολογιστές, εύρος ζώνης δικτύων, ή εξοπλισμός αποθήκευσης.

#### **4.2.2.5 Φορέας παροχής υπηρεσιών**

Ο φορέας παροχής υπηρεσιών είναι ή οργάνωση που παρέχει μια υπηρεσία οποιοδήποτε είδους.

#### **4.2.2.6 Ιδιοκτήτης λογισμικού**

Η οργάνωση που είναι κύρια των δικαιωμάτων πνευματικής ιδιοκτησίας στο λογισμικό. Οι ιδιοκτήτες του λογισμικού και μπορούν να χορηγήσουν άδεια σε άλλους (κάτοχοι άδειας) ώστε να χρησιμοποιηθεί το λογισμικό. Οι κάτοχοι άδειας δικαιώματα στο

λογισμικό όπως αυτά καθορίζονται από τη συμφωνία τους με τον ιδιοκτήτη του λογισμικού.

#### **4.2.2.7 Κάτοχος άδειας**

Αυτός είναι ο εξουσιοδοτημένος χρήστης του λογισμικού υπό τους όρους που χορηγούνται από έναν ιδιοκτήτη λογισμικού.

#### **4.2.2.8 Μεσίτης**

Αυτός είναι ο προξενητής ή αλλιώς μεσάζων (broker). Οι μεσίτες διαπραγματεύονται τις τιμές βασισμένες στις συναλλαγές με πολλούς προμηθευτές και πελάτες.

#### **4.2.2.9 Προμηθευτής δικτύων**

Ο προμηθευτής δικτύων είναι ο ιδιοκτήτης του δικτύου πάνω στο οποίο μεταφέρονται τα δεδομένα. Συχνά αυτό είναι ο χειριστής ενός δημόσιου δικτύου.

#### **4.2.2.10 Ανθρώπινος προμηθευτής πείρας**

Ο ανθρώπινος προμηθευτής πείρας είναι μία οντότητα που έχει τα πολύ καλά καταρτισμένα ανθρώπινα δυναμικά που μπορούν να προσθέσουν την αξία στις υπολογιστικές εργασίες. Ένα παράδειγμα είναι η υπηρεσία γνωμοδότησης προστιθέμενης αξίας σχετικά με κάποια υπηρεσία.

#### **4.2.2.11 Ελεγκτής**

Ο ελεγκτής (auditor) είναι μία οντότητα την οποία εμπιστεύονται τα συμβαλλόμενα μέρη μιας επιχειρηματικής σχέσης και εξετάζει κατά πόσο τηρείται η συμφωνία που έχουν.

#### **4.2.2.12 Ρυθμιστές**

Αυτοί είναι που καθορίζουν τους γενικούς κανόνες της αγοράς. Συνήθως αυτοί είναι οι κυβερνήσεις ή μεγαλύτερες οργανώσεις (όπως η Ευρωπαϊκή Επιτροπή) που καθορίζουν και επιβάλλουν την εθνική και διεθνή πολιτική.

#### **4.2.2.13 Πάροχος χρεώσεων**

Ο πάροχος χρεώσεων είναι μία οργάνωση που τιμολογεί τις υπηρεσίες για κάθε χρήστη με βάση τη χρήση τους.

#### **4.2.2.14 Ενορχηστρωτής**

Ένας ενορχηστρωτής έχει τα ρόλο της ενοποίησης και συγκέντρωσης υπηρεσιών οι οποίες προσφέρονται σε μια νέα υπηρεσία. Χαρακτηριστικό παράδειγμα είναι ένας πάροχος ολοκληρωμένων τηλεπικοινωνιών που προσφέρει σε ένα πακέτο σταθερή, κινητή τηλεφωνία και internet.

### **4.2.3 Επιχειρηματικές απαιτήσεις**

Αναλύοντας τα επιχειρηματικά μοντέλα προέκυψαν συγκεκριμένες απαιτήσεις προκειμένου τα πλέγματα και οι υπηρεσιοστρεφείς αρχιτεκτονικές να δημιουργήσουν τη δική τους δυναμική αγορά. Αυτές οι απαιτήσεις περιλαμβάνουν αναφέρονται σε συγκεκριμένες αρχές οι οποίες θα επιτρέψουν στα πλέγματα να είναι επιχειρηματικά βιώσιμα μέσα από την εκτέλεση σε αυτά επιχειρηματικών εφαρμογών καλύπτοντας πραγματικές ανάγκες. Ένα σύνολο τεχνικών προκλήσεων και αρχιτεκτονικών αρχών έχουν προκύψει από αυτές τις απαιτήσεις και έχουν επηρεάσει τον τρόπο σχεδιασμού και λειτουργίας των νέων αυτών υποδομών. Αυτές οι προκλήσεις αφορούν την ασφάλεια, τη διαχείριση, την απόδοση και τη χορήγηση αδειών (licensing). Οι αρχιτεκτονικές αρχές σχεδιασμού εξετάζουν τη δυναμική των συμφωνιών παροχής υπηρεσιών (Service Level Agreement - SLA) [2], τη δυναμική διαχείριση πόρων και τέλος των ελάχιστων στοιχείων που πρέπει να περιλαμβάνει η υποδομή προκειμένου να παρέχει την απαιτούμενη ποιότητα υπηρεσιών στους χρήστες.

Σε κάθε περίπτωση τα επιχειρηματικά μοντέλα είναι απαραίτητα για να ισορροπήσουν τους κινδύνους ενάντια στις ανταμοιβές. Για να είναι ένα επιχειρηματικό μοντέλο βιώσιμο, η ισορροπία των κινδύνων και των ανταμοιβών πρέπει να το καταστήσει

ευεργετικό για όλους τους συμμετέχοντες. Εάν το επιχειρηματικό μοντέλο είναι ασαφές ή δεν δίνει ικανοποιητικό πλεονέκτημα σε καθένα από τους συμμετέχοντές του, δεν θα είναι βιώσιμο. Σε αυτό το πλαίσιο οραματιζόμαστε μια παγκόσμια αγορά «πλέγματος» όπου οι πόροι, οι πληροφορίες και οι υπηρεσίες μπορούν να αγοραστούν και να πωληθούν. Κάτι τέτοιο φυσικά δεν έχει πραγματοποιηθεί ακόμα ωστόσο υπάρχουν αρκετά επιτυχημένα παραδείγματα μικρότερης κλίμακας που βρίσκονται σε εξέλιξη. Η δυνατότητα που αυτά προσφέρουν είναι να επιλεγτούν και να χρησιμοποιηθούν τμήματα υπηρεσιών από ποικίλες ανεξάρτητες πηγές και να ενσωματωθούν σε μια νέα εφαρμογή με νέες λειτουργίες.

Αυτό που υπόσχονται τα πλέγματα και οι υπηρεσιοστρεφείς αρχιτεκτονικές είναι να επιτρέψουν νέους τρόπους εργασίας καθώς και νέες εμπορικές ευκαιρίες, που στηρίζουν την ευελιξία και την ελαστικότητα, κάτι που γίνεται ακόμα πιο σημαντικό εάν οι επιχειρήσεις λειτουργούν ανταγωνιστικά σε αυτή την παγκόσμια αγορά. Για να γίνει όμως κάτι τέτοιο πρέπει η υποδομή να αναπτυχθεί υποστηρίζοντας την εμπορικά βιώσιμη χρήση των πόρων από τους οργανισμούς. Οι επιχειρησιακές σχέσεις κωδικοποιούνται γενικά στα όποια συμβόλαια συντάσσουν μεταξύ τους. Αυτά καθορίζουν ρητά όλες τις σχετικές λεπτομέρειες: τους εμπλεκόμενους, τις επιχειρηματικές πρακτικές, τα πρότυπα που χρησιμοποιούνται, καθώς επίσης και την τιμολόγηση και τις ποινικές ρήτρες για την παραβίαση της συμφωνίας. Στον κόσμο των υπηρεσιών και των πλεγμάτων αυτό μεταφράζεται σε συμφωνία επιπέδων εξυπηρέτησης (SLA). Ένα SLA μπορεί να χρησιμοποιηθεί για να παρέχει ένα ρητό πλαίσιο για τις σχέσεις μεταξύ των οντοτήτων πλέγματος (πόροι, άτομα, ή οργανισμοί). Με βάση αυτές τις εκτιμήσεις μπορούμε να προσδιορίσουμε τις ακόλουθες οκτώ απαιτήσεις που μια βιώσιμη υποδομή πλέγματος ή υπηρεσιοστρεφής υποδομή πρέπει να ικανοποιήσει:

#### **4.2.3.1 Ευέλικτα επιχειρηματικά μοντέλα**

Τα πλέγματα και οι υπηρεσιοστρεφείς υποδομές μπορούν πιθανώς να υποκινήσουν νέα επιχειρηματικά μοντέλα όμως αυτό που χρειάζεται κυρίως είναι να μπορέσουν να υποστηρίξουν βιώσιμα επιχειρηματικά μοντέλα. Όπως με το διαδίκτυο, όπου διάφορα επιχειρηματικά μοντέλα λειτουργούν, τα πλέγματα πρέπει να είναι προσαρμόσιμα σε ποικίλα πρότυπα. Οι νέες υποδομές πρέπει να προσφέρουν τα πραγματικά οφέλη σε μια σειρά από νέες επιχειρηματικές εφαρμογές. Συγκεκριμένα, πρέπει να υπάρξουν οικονομικά πλεονεκτήματα στην προσέγγιση πλέγματος πέρα από τις συμβατικές λύσεις που επικεντρώνονται συνήθως σε τεχνολογικά και ερευνητικά θέματα.

#### **4.2.3.2 Συγκεκριμένοι όροι παροχής ποιότητας υπηρεσιών**

Η λειτουργική περιγραφή μιας υπηρεσίας, που χρησιμοποιείται για τη διαφήμιση ή ως τμήμα του SLA, δεν καλύπτει όλες τις ανάγκες ενός καταναλωτή ή ενός παρόχου μιας υπηρεσίας. Προκειμένου να επιτευχθεί η διαφοροποίηση μεταξύ των προμηθευτών, όροι ποιότητας εξυπηρέτησης (QoS) όπως η αξιοπιστία, η ευρωστία, και η ανθεκτικότητα είναι απαραίτητοι για την αξιολόγηση της παρεχόμενης υπηρεσίας. Ο αριθμός αυτών των όρων ποιότητας υπηρεσίας είναι απροσδιόριστος και καθορίζονται από την ίδια την εφαρμογή και την υπηρεσία. Έτσι τα SLAs πρέπει να υποστηρίξουν τους όρους ελαστικά. Επιπλέον, για να υποστηρίξουν τη διαχείριση της υποδομής και από πλευράς καταναλωτών και προμηθευτών, οι συμφωνίες πρέπει να είναι ευέλικτες ως προς τους όρους που περιλαμβάνουν. Τέλος, οι μέθοδοι με τις οποίες θα χρησιμοποιούνται τα SLAs πρέπει να είναι τυποποιημένες, κατάλληλες για αυτοματοποίηση τόσο για τους καταναλωτές όσο και τους παροχείς υπηρεσιών.

#### **4.2.3.3 Ασφάλεια**

Σε όλα σχεδόν τα επιχειρηματικά μοντέλα η ασφάλεια είναι πολύ σημαντική. Κατά συνέπεια η λειτουργική ασφάλεια και η λειτουργική ακεραιότητα πρέπει να έχουν μεγάλη

προτεραιότητα και στα υπολογιστικά περιβάλλοντα. Εκτός από τις κοινές ανησυχίες για ασφάλεια επικοινωνιών (επικύρωση, έγκριση, ακεραιότητα και εμπιστευτικότητα) και λειτουργική ασφάλεια (ανίχνευση παρείσφρησης, λογιστική, αξιολόγηση του κινδύνου, και λογιστικός έλεγχος), υπάρχει μια απαίτηση για δυναμικά μεταβαλλόμενες σχέσεις εμπιστοσύνης. Αυτές οι αλλαγές σε σχέσεις εμπιστοσύνης πρέπει να ρυθμιστούν σε όλη τη διάρκεια ζωής μιας επιχειρηματικής σχέσης που ταυτίζεται με τη διάρκεια ζωής του SLA.

#### **4.2.3.4 Δυναμική Σύνθεση**

Ιδιαίτερα σημαντική θεωρείται η δυνατότητα δημιουργίας νέων πιο ειδικευμένων υπηρεσιών. Αυτό περιλαμβάνει τη σύνθεση υπηρεσιών κατά τη διάρκεια ζωής της επιχειρηματικής σχέσης, κυρίως κατά τη διάρκεια της εκτέλεσης, και απαιτεί τα συνθετικά στοιχεία-υπηρεσίες, οι σημασιολογικές περιγραφές των υπηρεσιών, και οι τεχνολογίες ροής εργασιών (workflows) να λειτουργούν δυναμικά και συντονισμένα με άξονα πάντα την ασφάλεια όπως αυτή καθορίζεται στη συμφωνία.

#### **4.2.3.5 Οικονομική βιωσιμότητα**

Η βιώσιμη επέκταση των πλεγμάτων σε ένα εμπορικό περιβάλλον απαιτεί την ισχυρή οικονομική υποστήριξη. Το εμπόδιο αυτό αποτρέψτε μέχρι σήμερα τις υποδομές πλέγματος από το να γίνουν οικονομικά βιώσιμες όπως το διαδίκτυο. Η υποστήριξη για την ανταγωνιστική προμήθεια των υπηρεσιών είναι ουσιαστική στο να επιτρέψει τη μετάβαση σε μια ανοικτή αγορά.

#### **4.2.3.6 Ιδιωτικότητα**

Η ιδιωτικότητα, που είναι ουσιαστικά η εφαρμογή και η διαχείριση της εμπιστευτικότητας σύμφωνα με τις πολιτικές, τις προτιμήσεις και τους νόμους, πρέπει να είναι σε μια υποδομή πλέγματος μια κεντρική πτυχή.



#### **4.2.3.7 Αποδοτική και ευέλικτη διαχείριση**

Η δυνατότητα να διαχειρίζεται μια υποδομή πλέγματος αποτελεσματικά και αποδοτικά είναι βασική για την εμπορική βιωσιμότητά της. Η χειρωνακτική διαχείριση, ιδιαίτερα μιας μεγάλης δυναμικής υποδομής, είναι ακριβή και περιορίζει τον αριθμό βιώσιμων επιχειρηματικών μοντέλων. Η δυνατότητα να μπορεί μια υποδομή να διαχειρίζεται και να λειτουργεί με ελάχιστο κόστος είναι μια σημαντική απαίτηση. Οι βιώσιμες υποδομές πρέπει να λειτουργούν με έναν υψηλό βαθμό της αυτοματοποίησης επειδή οι λειτουργικές δαπάνες της υποδομής καθορίζονται σε μεγάλο βαθμό από τις δαπάνες για προσωπικό. Επιπλέον, οποιαδήποτε διοικητική υποδομή πρέπει να ενσωματώσει στη λειτουργία της τη διαχείριση των SLAs.

#### **4.2.3.8 Διαδραστική υποστήριξη**

Η παροχή διαδραστικής υποστήριξης μπορεί να είναι αναλύεται σε διάφορες ανεξάρτητες ιδιότητες μιας σχέσης υπηρεσιών, π.χ. ο χρόνος απόκρισης, έλεγχος, και ιδιωτικότητα. Οι εφαρμογές πλέγματος πρέπει να παρέχουν μηχανισμούς που να μπορούν να υποστηρίξουν τις επεκτάσεις υπηρεσιών. Αυτό ασκεί επίδραση τόσο στο SLA όσο και στη βασική υποδομή πλέγματος.

### **4.3 Αρχές σχεδιασμού της αρχιτεκτονικής**

Το κεφάλαιο αυτό περιγράφει τις βασικές αρχές σχεδιασμού της αρχιτεκτονικής πλεγμάτων και υπηρεσιοστρεφών αρχιτεκτονικών. Αυτές οι αρχές προέρχονται από τις προκλήσεις που περιγράφηκαν αναλυτικά στα προηγούμενα κεφάλαια καθώς και από την εξέταση των επιχειρηματικών μοντέλων, ενώ αναπτύχθηκαν στο πλαίσιο του ευρωπαϊκού ερευνητικού έργου NextGRID [50][68]. Τις αρχές θα μπορέσαμε να τις κατηγοριοποιήσουμε σε κύριες και δευτερεύουσες.

Οι κύριες αρχές είναι:

- **Σχεδιασμός με βάση τις συμφωνίες επιπέδου υπηρεσιών (SLAs):** Τα SLAs πρέπει να είναι βασικές δομικές μονάδες στα πλέγματα και τις υπηρεσιοστρεφείς υποδομές ενώ η δυναμική συμπεριφορά τους είναι σημαντική στη λειτουργία και διαχείρισή τους.
- **Δυναμική σύνθεση υπηρεσιών:** Η δυναμική σύνθεση των υπηρεσιών και πόρων είναι ένας παράγοντας κλειδί στην καθιέρωση λειτουργικών επιχειρησιακών υποδομών.
- **Ελάχιστη υποδομή:** Όλες οι υποδομές πρέπει να είναι απλές και να εξασφαλίζουν ευκολία της συντήρησης και μεγάλο εύρος δυνατοτήτων. Ωστόσο πρέπει να έχει ικανοποιητικά χαρακτηριστικά γνωρίσματα ώστε να υποστηρίξει βιώσιμα επιχειρηματικά μοντέλα.

Οι δευτερεύουσες αρχιτεκτονικές αρχές μπορούν να βελτιώσουν την ανάπτυξη, την επέκταση, και τη διαχείριση του πλέγματος καθιστώντας το πιο αποδοτικό:

- **Δυναμική διάρκεια ζωής υπηρεσιών:** Οι υπηρεσίες πρέπει να υπάρχουν διαθέσιμες όταν απαιτούνται, αλλά εξαφανίζονται όταν πλέον δεν απαιτούνται.
- **Δυναμικό περιεχόμενο:** Το περιεχόμενο των υπηρεσιών πρέπει να είναι σε θέση να αυξηθεί και εξελιχθεί κατά τη διάρκεια της διάρκειας ζωής μιας υπηρεσίας.
- **Αυτόνομη Διαχείριση:** Τα πλέγματα πρέπει να είναι τόσο εύχρηστα όσο και αυτόνομα.
- **Αναζήτηση Υπηρεσιών:** Οποιοδήποτε πλέγμα πρέπει να επιτρέπει την αναζήτηση των υπηρεσιών.
- **Ανοικτό σχεδιασμό και αναπτυξιακή διαδικασία:** Οι νέες υποδομές θα πρέπει να είναι ιδιαίτερα κατανεμημένες και να αποτελούνται από υπηρεσίες που

προσφέρονται από διαφορετικούς παρόχους. Κατά συνέπεια τα δομικά στοιχεία πρέπει να είναι διαλειτουργικά τόσο κατά το σχεδιασμό όσο και κατά τη λειτουργία της υποδομής.

### **4.3.1 Κύριες αρχές σχεδιασμού**

#### **4.3.1.1 Σχεδιασμός με βάση τις συμφωνίες επιπέδου υπηρεσιών (SLAs)**

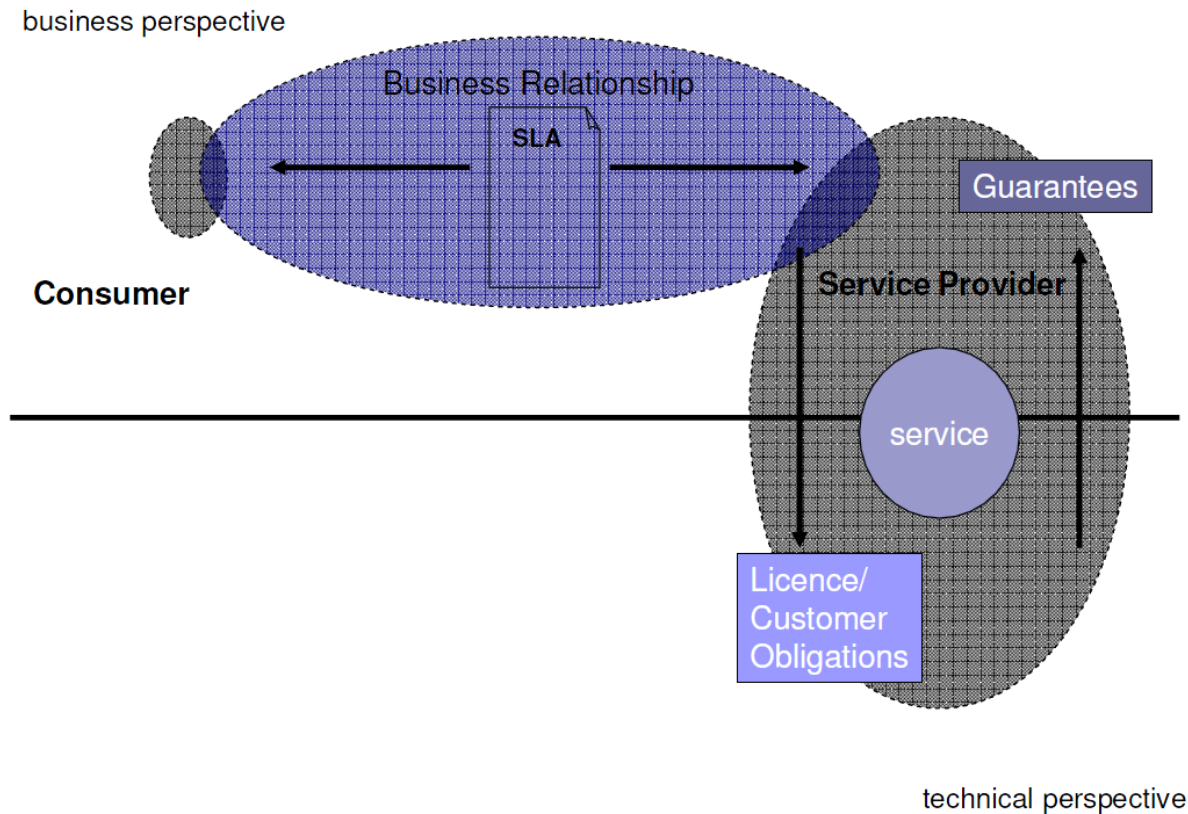
Η συμφωνία σε επίπεδο υπηρεσιών (SLA) είναι μείζονος σημασίας όπως είδαμε στα προηγούμενα επιχειρηματικά μοντέλα αφού καθορίζει το επίπεδο ποιότητας υπηρεσιών. Κατά συνέπεια αναμένεται να έχει κεντρικό ρόλο στην αρχιτεκτονική του πλέγματος ή της υπηρεσιοστρεφής υποδομής. Όλες οι αλληλεπιδράσεις μεταξύ των δομικών συστατικών και κάθε οντότητας περιγράφονται μέσω του SLA επιτρέποντας την παρακολούθηση της εκτέλεσης της εργασίας, τη διαχείριση της υποδομής και τον έλεγχο σε λογιστικό επίπεδο (accounting/billing). Η αρχιτεκτονική προσέγγιση λοιπόν βασισμένη στα SLAs παρέχει ένα ομοιόμορφο πλαίσιο για τη διαχείριση και τη λειτουργία όλων των πτυχών διασφάλισης ποιότητας υπηρεσιών στην υποδομή όπως η απόδοση, η ασφάλεια, η διαχείριση πολιτικών, τους κανονισμούς ιδιωτικότητας, κ.λπ.

Ένα SLA καθορίζει τη φύση και τις συνέπειες μιας αλληλεπίδρασης μεταξύ ενός παρόχου υπηρεσιών και ενός καταναλωτή. Η συμφωνία γίνεται σε κάποιο επιχειρηματικό πλαίσιο, και περιλαμβάνει τις αποφάσεις που λαμβάνονται από κάθε συμβαλλόμενο μέρος σε μέρος της συμφωνίας και καθιστούν τις όρους συμφωνίας αποδεκτούς και από τις δύο πλευρές. Ένα SLA δημιουργείται πριν την εκτέλεση της υπηρεσίας και βρίσκεται σε ισχύ καλύπτοντας ολόκληρο τον κύκλο ζωής της. Εντούτοις, είναι δυνατό να υπάρξουν SLAs με μια υπάρχουσα υπηρεσία, π.χ. μέσω μιας διαδικασίας ενορχήστρωσης υπηρεσιών από έναν υπάρχοντα καταναλωτή, παράγοντας νέες αλληλεπιδράσεις με άλλους καταναλωτές.

Το SLAs επομένως έχει μια τεράστια επιρροή σε όλες τις πτυχές της υπηρεσίας και της υποδομής πάνω στην οποία η υπηρεσία εκτελείται. Πριν από τη δημιουργία ενός SLA,

κάθε συμβαλλόμενο μέρος αναπτύσσει δραστηριότητες μέσα σε ένα ιδιωτικό πλαίσιο. Μετά από τη δημιουργία ενός SLA, ένα κοινό πλαίσιο καθιερώνεται και όλο το περιεχόμενο αυτού του κοινού πλαισίου ενσωματώνεται (άμεσα ή έμμεσα) στο SLA. Να σημειώσουμε ότι στη δημιουργία ενός SLA, τα στοιχεία του ιδιωτικού πλαισίου των δύο συμβαλλόμενων μερών πρέπει να ευθυγραμμίσουν για να επιτρέψουν στα δύο συμβαλλόμενα μέρη για να δημιουργήσουν το SLA και ως εκ τούτου να καθιερώσουν ένα κοινό πλαίσιο για την υπηρεσία. Επιπρόσθετα, ενώ πολλές πτυχές του πλαισίου, μετά από τη δημιουργία ενός SLA, μπορούν να μοιραστούν ή να συμφωνηθούν (π.χ. το επίπεδο ποιότητας υπηρεσιών), τα στοιχεία λειτουργίας που δημιουργούν το κοινό πλαίσιο μπορεί να παραμένουν κρυφά. Για παράδειγμα κανένας εμπορικός φορέας παροχής υπηρεσιών δεν θα αποκαλύψει το σχέδιο των πόρων του και την ακριβή τους ποσότητα σε έναν καταναλωτή, απλά θα συμφωνήσει να προσφέρει μια υπηρεσία με την ποιότητα υπηρεσίας και τους πόρους που ζητήθηκαν.

Ένα SLA υφίσταται μεταξύ δύο συμβαλλόμενων μερών, του φορέα παροχής υπηρεσιών και του καταναλωτή. Ωστόσο σε μερικές περιπτώσεις υπάρχει ένας επιπλέον ρόλος. Ο ρόλος αυτός μπορεί να περιλαμβάνει μια ανεξάρτητης αρχή επαλήθευσης και ελέγχου στην περίπτωση των διαφωνιών μεταξύ καταναλωτή και παρόχου. Η δομή λοιπόν του SLA πρέπει να μπορεί να παρέχει όλες τις πληροφορίες που απαιτούνται ακολουθώντας συγκεκριμένα πρότυπα και δομή που θα επιτρέπει αυτοματοποιημένη, και επομένως οικονομικότερη, επεξεργασία του. Στο πλαίσιο αυτό οι όροι ενός SLA μπορούν να ταξινομηθούν σε «δυναμικούς» και «στατικούς». Οι στατικοί οροί περιλαμβάνουν περισσότερο οικονομικά στοιχεία, τα εμπλεκόμενα μέρη και τη διάρκεια ζωής του. Στους δυναμικούς όρους περιλαμβάνονται οι όροι υψηλού επιπέδου, οι οποίοι γίνονται περισσότερο κατανοητοί από τους καταναλωτές ή τις εφαρμογές.



**Σχήμα 9: Η αντιστοιχία των απαιτήσεων των καταναλωτών με τους πόρους των υπηρεσιών μέσω της χρήσης συμφωνιών επιπέδου υπηρεσιών**

Διαπιστώνουμε λοιπόν την ανάγκη (όπως περιγράφεται στο προηγούμενο σχήμα) αντιστοιχίας των απαιτήσεων των καταναλωτών, σε επιχειρηματικό επίπεδο για μια υπηρεσία με τους πόρους που χρειάζονται από τον πάροχο προκειμένου να την προσφέρει. Συγκεκριμένα απαιτείται:

- αντιστοίχιση των απαιτήσεων από επιχειρηματικό επίπεδο και περιλαμβάνονται στο SLA σε πολιτικές λειτουργίας στην υποδομή του παρόχου και σε συγκεκριμένες τεχνικές παραμέτρους για τους πόρους.
- αντιστοίχιση των πληροφοριών που προέρχονται από την παρακολούθηση λειτουργίας των πόρων σε τεχνικό επίπεδο με τους όρους που περιλαμβάνονται στο SLA για να ελέγχεται η σωστή λειτουργία της υπηρεσίας και το επίπεδο ποιότητας με το οποίο προσφέρεται.

*Η σελίδα αυτή είναι σκόπιμα λευκή*

Πρέπει να σημειώσουμε ότι αυτή η αντιστοίχιση μεταξύ επιχειρηματικού και τεχνικού επιπέδου δεν αφορά μόνο τις επιδόσεις και την ποιότητα υπηρεσιών. Σε ορισμένες περιπτώσεις μπορεί να είναι απαραίτητο να χρησιμοποιηθούν άδειες χρήσης για το λογισμικό προκειμένου να εκπληρωθεί το SLA που σημαίνει ότι ο πάροχος θα πρέπει να προβεί σε ανάλογες ενέργειες για να εξασφαλίσει τις άδειες. Αν όμως το SLA προβλέπει ότι ο χρήστης θα παρέχει τις άδειες, θα πρέπει να γίνουν από την πλευρά του παρόχου ενέργειες σε επίπεδο ασφάλειας προκειμένου να μπορούν να χρησιμοποιηθούν οι άδειες.

#### **4.3.1.2 Δυναμική σύνθεση υπηρεσιών**

Στα σύγχρονα πλέγματα υπολογιστών και υπηρεσιοστρεφών υποδομών, οι εφαρμογές θα κατασκευαστούν με τη σύνθεση διάφορων υπηρεσιών. Μια βασική απαίτηση είναι ότι μηχανισμοί σύνθεσης θα πρέπει να οδηγούν σε συγκεκριμένες δομές με κοινά χαρακτηριστικά ακολουθώντας αυστηρούς κανόνες σύνθεσης. Το κλειδί σε αυτό είναι πάλι τα SLAs τα οποία καθορίζουν όλες τις αλληλεπιδράσεις μεταξύ των συστατικών μερών. Εδώ παρουσιάζουμε κανόνες σύνθεσης που μπορούν να εξασφαλίσουν τη σύνθεση υπηρεσιών με δυναμικό και αυτόματο τρόπο. Από την ανάλυση προέκυψαν τρεις βασικοί κανόνες σύνθεσης που αναλύονται παρακάτω:

**Η κοινή εκμετάλλευση πόρων** είναι η περίπτωση όπου ένας καταναλωτής μιας υπηρεσίας την μοιράζεται με έναν άλλο καταναλωτή:

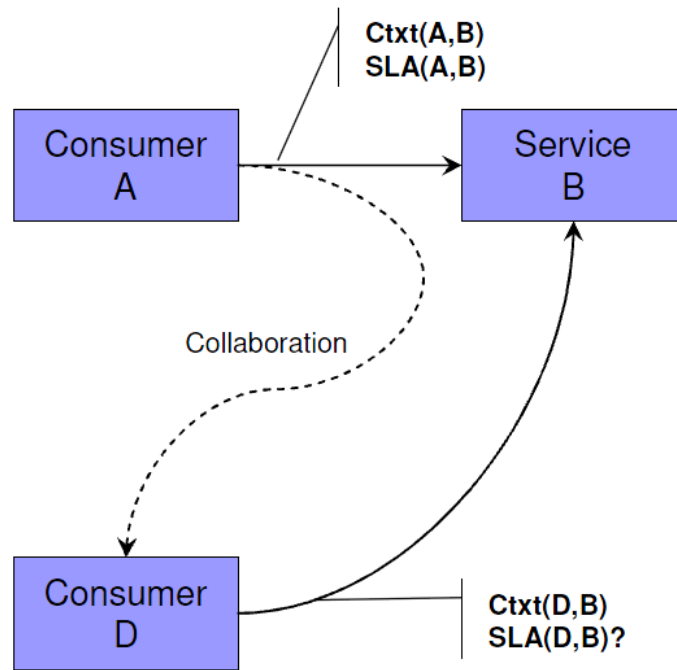
Η κοινή εκμετάλλευση πόρων είναι ουσιαστικά ομοσπονδία μεταξύ καταναλωτών οι οποίοι αλληλεπιδρούν με το φορέα παροχής υπηρεσιών. Η κοινή εκμετάλλευση πόρων είναι πολύ σημαντική και αρκετά διαδεδομένη στα πλέγματα με επιχειρηματικό προσανατολισμό.

**Ο σύνθεση πόρων** προκύπτει όταν ένας καταναλωτής δύο υπηρεσιών ζητά από τις υπηρεσίες να αλληλεπιδράσουν με συγκεκριμένο τρόπο.

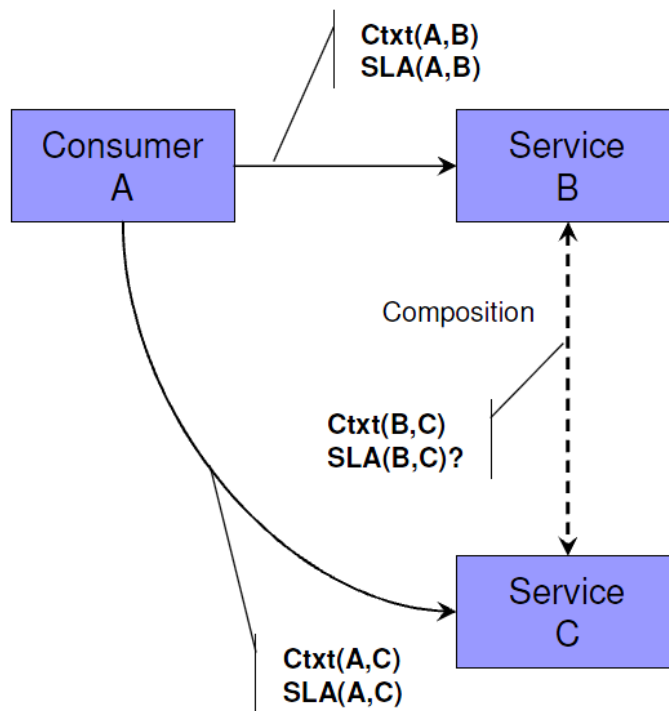
Αυτή η διαδικασία συνδυάζει αποτελεσματικά τους πόρους από δύο φορείς παροχής υπηρεσιών για να ικανοποιήσει τις ανάγκες του καταναλωτή. Να σημειωθεί ότι είναι εφικτό να αυτοματοποιηθεί το πλαίσιο αλληλεπίδρασης των υπηρεσιών βασισμένο στους όρους αλληλεπιδράσεις που περιλαμβάνονται στα αρχικά SLA.

**Η ενθυλάκωση των πόρων** προκύπτει όταν ένας φορέας παροχής υπηρεσιών προσφέρει μια υπηρεσία χρησιμοποιώντας μιας άλλης υπηρεσία, χωρίς την άμεση αλληλεπίδραση μεταξύ του παρόχου της δεύτερης υπηρεσίας και του καταναλωτή της πρώτης. Η ενθυλάκωση μπορεί να αρχίσει και από τις δύο πλευρές. Η ενδιάμεση υπηρεσία (B) μπορεί χρησιμοποιείται από τον καταναλωτή (A) και να αποφασίσει κάποια στιγμή να αναθέσει μέρος της εργασίας σε μία υπηρεσία (C) ή από την άλλη η μεσάζουσα υπηρεσία να είχε καθιερώσει τη σχέση της με την υπηρεσία (C) πριν από την προσφορά υπηρεσιών στον καταναλωτή (A). Να σημειώσουμε ότι σε κάθε περίπτωση ο (B) δεν είναι υποχρεωμένος να γνωστοποιήσει τις λεπτομέρειες της αλληλεπίδρασής του με τον (C) στον (A) και το αντίθετο.

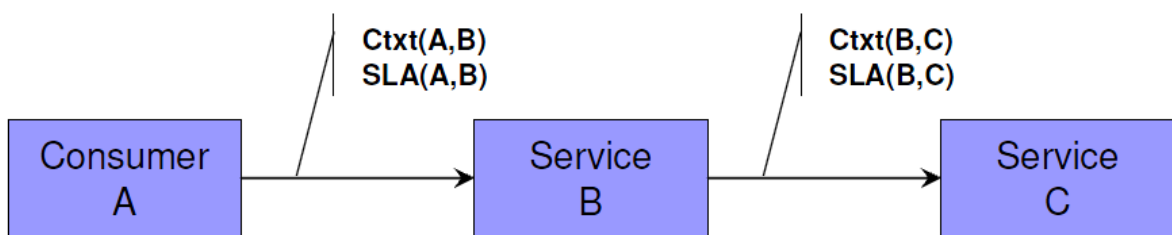




Σχήμα 10: Κοινή εκμετάλλευση πόρων



Σχήμα 11: Σύνθεση πόρων



Σχήμα 12: Ενθολάκωση πόρων

*Η σελίδα αυτή είναι σκόπιμα λευκή*

Σε κάθε μία από τις παραπάνω περιπτώσεις ο ρόλος των SLAs είναι καθοριστικός, ενώ μπορεί να υφίστανται σε διάφορα επίπεδα και με διάφορους τρόπους ανάλογα πάντα με το επιχειρηματικό μοντέλο και την εφαρμογή. Κατά συνέπεια η ποιότητα υπηρεσιών όπως καθορίζεται από τους όρους των SLAs βρίσκεται και αυτή σε διάφορα επίπεδα της υποδομής και εκφράζεται με διαφορετικό τρόπο. Επομένως για να είναι η συνολική υποδομή λειτουργική και το επιχειρηματικό μοντέλο εφικτό θα πρέπει οι οντότητες που βρίσκονται σε κάθε επίπεδο να «μιλούν την ίδια γλώσσα» δηλαδή να μπορούν να καταλάβουν τη σημασία των όρων των SLAs και της ποιότητας υπηρεσίας, ενώ θα πρέπει να είναι και εφικτή η «μετάφραση» των παραπάνω από το ένα επίπεδο στο άλλο.

#### ***4.3.1.3 Ελάχιστη Υποδομή***

Η αρχή της ελάχιστης υποδομής ανάγεται στο ότι η επικοινωνία, η συμπεριφορά και οι υπηρεσίες θα γίνονται κάτω από συγκεκριμένα πρότυπα τα οποία προβλέπουν την απλούστερη / βασική υποδομή. Η υποδομή αυτή θα μπορεί να προσαρμοστεί και να επεκταθεί με επιπλέον δυνατότητες και χαρακτηριστικά για να ικανοποιήσει το επιχειρηματικό μοντέλο.

#### ***4.3.2 Δευτερεύουσες αρχές σχεδιασμού***

Οι δευτερεύουσες αρχιτεκτονικές αρχές εμπίπτουν σε πέντε κατηγορίες.

##### ***4.3.2.1 Δυναμική διάρκεια ζωής υπηρεσιών***

Οι υπηρεσίες σε περιβάλλοντα πλεγμάτων και υπηρεσιοστρεφών υποδομών είναι ιδιαίτερα κατανεμημένα. Οι χρήστες και οι πάροχοι πρέπει να διαχειρίζονται τη διάρκεια ζωής των υπηρεσιών για να εξασφαλιστεί η διαθεσιμότητά τους όταν αυτό απαιτείται, αλλά δεν καταναλώνουν πόρους όταν αυτό δεν χρειάζεται. Η διάρκεια ζωής των υπηρεσιών διαπραγματεύεται μεταξύ των χρηστών και των προμηθευτών τους. Αυτό πρέπει να είναι δυναμικό καθώς οι περιστάσεις αλλάζουν κατά τη διάρκεια της διάρκειας

ζωής μιας υπηρεσίας. Η καταστροφή των υπηρεσιών πρέπει να είναι ρητά ορισμένη, μόλις τελειώσει η χρήση τους από τους καταναλωτές.

#### **4.3.2.2 Δυναμικό περιεχόμενο**

Οι υπηρεσίες στα νέα αυτά περιβάλλοντα θα έχουν διαφορετικά χαρακτηριστικά αλλά ενδέχεται να εξυπηρετούνται (χρησιμοποιούν πόρους) από την ίδια υποδομή. Επίσης το περιεχόμενο της ίδιας της υπηρεσίας μπορεί να μεταβάλλεται δυναμικά κατά τη διάρκεια της διάρκειας ζωής της υπηρεσίας. Κατά συνέπεια, οι γλώσσες που περιγράφουν αυτό το περιεχόμενο πρέπει να είναι κατάλληλες για να το περιγράψουν και φυσικά το ίδιο ισχύει και για την υπηρεσία και το επίπεδο ποιότητάς της.

#### **4.3.2.3 Αυτόνομη διαχείριση**

Τα δυναμικά πλέγματα πρέπει να μπορούν να διαχειριστούν εύκολα. Τα μεγάλα, κατανεμημένα πλέγματα έχουν μεγάλες απαιτήσεις και είναι εφικτά μόνο εάν μπορούν να διαχειριστούν αυτόνομα. Ο τελικός σκοπός για τα πλέγματα και γενικά τις υπηρεσιοστρεφείς υποδομές είναι ότι πρέπει είναι ικανά για αυτόνομη διαχείριση, ακολουθώντας συγκεκριμένες πολιτικές όπως ορίζουν οι ιδιοκτήτες τους.

#### **4.3.2.4 Αναζήτηση υπηρεσιών**

Η αρχιτεκτονική των νέων υποδομών θα πρέπει να υποστηρίζει εκτεταμένες μεθόδους αναζήτησης υπηρεσιών μέσω μητρώων που θα χρησιμοποιούνται ως συλλέκτες πληροφοριών.

#### **4.3.2.5 Ανοικτό σχεδιασμό και αναπτυξιακή διαδικασία**

Οι υπηρεσιοστρεφείς υποδομές είναι αρκετά κατανεμημένες και αποτελούνται από υπηρεσίες που αναπτύσσονται και από έναν μεγάλο αριθμό διαφορετικών ανθρώπων. Ένα τέτοιο οικοσύστημα θα μπορεί να λειτουργήσει μόνο εάν όλα τα συστατικά είναι διαλειτουργικά μέσα από εύκολες και απλές διαδικασίες. Οποιαδήποτε μη-διαλειτουργικότητα θα δημιουργούσε υποδομές που δεν μπορούν να επικοινωνήσουν

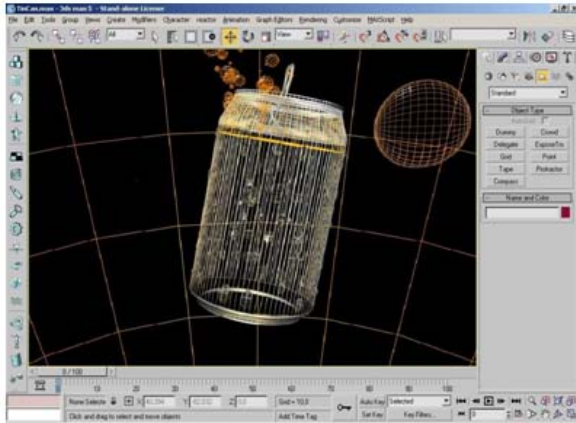
μεταξύ τους. Η χρήση κοινών αρχιτεκτονικών αρχών και ανοιχτών προτύπων βοηθά προς τη σωστή κατεύθυνση.

#### ***4.4 Εφαρμογή σε ένα καινοτόμο σενάριο χρήσης: Υπηρεσίες πολυμέσων σε περιβάλλον πλέγματος***

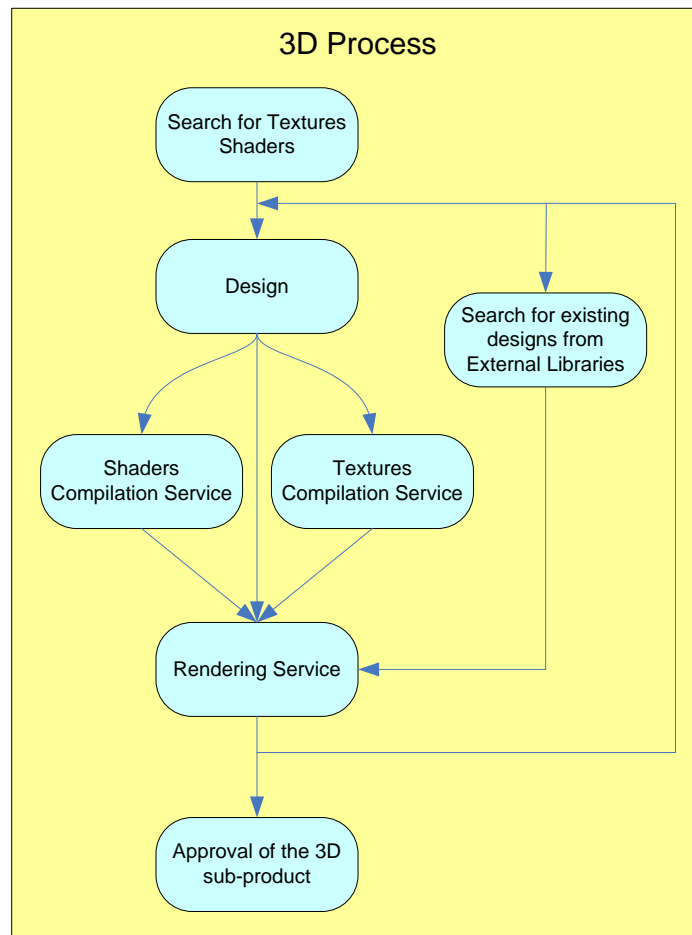
Το ερώτημα που τίθεται πριν την προσαρμογή και ένταξη μιας εφαρμογής σε ένα πλέγμα είναι γιατί να γίνει αυτό και να μην επιλεγθούν τοπικές λύσεις συστοιχίας υπολογιστών (clusters). Σε κάθε εφαρμογή η απάντηση είναι διαφορετική αλλά σε κάθε περίπτωση αυτό που παίζει πάντα ρόλο είναι το επιχειρηματικό μοντέλο που έχει αναπτυχθεί γύρω από την εφαρμογή και την ένταξη της στο πλέγμα. Για τους ερευνητικούς σκοπούς της διατριβής μελετήσαμε τις εφαρμογές πολυμέσων. Από μια γρήγορη ανάλυση για τις εφαρμογές πολυμέσων διαπιστώνουμε ότι τόσο ο αριθμός των χρηστών όσο και το είδος τους ποικίλει. Υπάρχουν σχεδιαστές ή μικρές εταιρίες, οι οποίοι εργάζονται ανεξάρτητα και ο μόνος τρόπος να βρεθούν οι υπολογιστικοί (ή άλλοι) πόροι είναι να γίνουν οι συμφωνίες με εξωτερικούς συνεργάτες και να πληρώσουν για τη χρήση των πόρων. Το περιβάλλον πλέγματος φαίνεται η ιδανική λύση για να καλύψει τις ανάγκες τους. Επιπλέον οι επιχειρήσεις που έκαναν τις επενδύσεις σε υποδομές βλέπουν τους πόρους τους να χρησιμοποιούνται λιγότερο από όσο προβλέπεται στο επιχειρηματικό τους σχέδιο ενώ άλλοι χρειάζονται πρόσθετους πόρους ώστε να εξασφαλίσουν την υπολογιστική ισχύ που χρειάζονται σε περιόδους αυξημένου φόρτου εργασίας. Εύκολα καταλαβαίνουμε λοιπόν ότι κάτω από αυτές τις περιστάσεις η επένδυσή τους δεν ικανοποιεί τις προσδοκίες τους.

Μια από τις πιο σημαντικές διαδικασίες στις εφαρμογές πολυμέσων είναι η τρισδιάστατη απεικόνιση (3D animation) ενώ παράλληλα παρουσιάζει μεγάλο ενδιαφέρον και ιδιαίτερα χαρακτηριστικά για ένταξη σε περιβάλλον πλέγματος. Η διαδικασία της τρισδιάστατης απεικόνισης μπορεί να είναι αποδοτικότερη μέσω της εισαγωγής ενός νέου

επιχειρηματικού μοντέλου. Αυτό το νέο επιχειρηματικό μοντέλο μπορεί να υποστηριχθεί από τις τεχνολογίες πλέγματος το οποίο μπορεί να διαχειριστεί την παραγωγή κάθε λογής εργασιών τρισδιάστατης απεικόνισης. Στις επόμενες εικόνες περιγράφεται η διαδικασία της τρισδιάστατης απεικόνισης (μετατροπή από wireframe στην εξομαλυμένη εικόνα – rendered image) καθώς και τα επιμέρους βήματα που απαιτούνται.



Σχήμα 13: Ο σχεδιασμός και το αποτέλεσμα τρισδιάστατων απεικονίσεων



Σχήμα 14: Η διαδικασία δημιουργίας τρισδιάστατων απεικονίσεων

*Η σελίδα αυτή είναι σκόπιμα λευκή*



#### **4.4.1 Ποιότητα υπηρεσιών και επιχειρηματικά μοντέλα για εφαρμογές πολυμέσων σε περιβάλλον πλέγματος**

Για να μπορέσουμε τελικά να ερευνήσουμε πως θα διασφαλίσουμε την ποιότητα υπηρεσιών σε μια εφαρμογή πολυμέσων πρέπει πρώτα να την προσδιορίσουμε. Σύμφωνα με όσα αναφέραμε στα προηγούμενα κεφάλαια η ποιότητα υπηρεσιών έχει άμεση σχέση με το επιχειρηματικό μοντέλο και πως αυτό εφαρμόζεται μέσα από τη χρήση νέων τεχνολογιών πλέγματος. Για παράδειγμα πρέπει να γνωρίζουμε τους τελικούς χρήστες, τι ακριβώς επιδιώκουν, με ποιο τρόπο αλληλεπιδρούν με τους παρόχους. Στο πλαίσιο λοιπόν που περιγράψαμε προηγούμενα αναλύουμε το επιχειρηματικό μοντέλο και στη συνέχεια την ποιότητα υπηρεσιών και τις αρχιτεκτονικές προκλήσεις που προκύπτουν. Επιπρόσθετα οι πληροφορίες αυτές θα χρησιμοποιηθούν σε επόμενο κεφάλαιο για τη δημιουργία σεναρίων, στα οποία βασίστηκαν τα πειράματα που εκτελέστηκαν.

##### **4.4.1.1 Ποια είναι η προσφορά της υπηρεσίας στον πελάτη;**

Σε ένα περιβάλλον πλέγματος οι πάροχοι εφαρμογών πολυμέσων θα μπορούσαν να προσφέρουν τις υπηρεσίες τους με διαφόρους τρόπους. Έτσι μπορούν να υπάρχουν υπηρεσίες που είναι ακριβές ή φτηνές, συνηθισμένες ή με ιδιαίτερα χαρακτηριστικά και λειτουργίες, γρήγορες ή «βαριές» σε υπολογιστικό επίπεδο, και φυσικά όλους τους συνδυασμούς των ανωτέρω. Αυτός σημαίνει ότι η επιχείρηση παραγωγής (πολυμέσων) επιλέγει από έναν κατάλογο υπηρεσιών αυτή που ικανοποιεί τις περισσότερες από τις απαιτήσεις της συγκεκριμένης εργασίας. Συνεπώς, οι σχεδιαστές έχουν πρόσβαση στις υπηρεσίες που ενδεχομένως δεν είναι διαθέσιμες εσωτερικά στην επιχείρηση και εκτελούν τις εργασίες τους σε ένα περιβάλλον με τους απεριόριστους (θεωρητικά) υπολογιστικούς πόρους. Επιπρόσθετα το πλέγμα μπορεί να προσφέρει υπηρεσίες που δεν θα μπορούσαν να είναι διαθέσιμες στους χρήστες είτε λόγω του κόστους τους (εφαρμογή και άδειες) ή λόγω των απαιτήσεων σε υλικό εξοπλισμό. Αυτές οι υπηρεσίες μπορεί να συμπληρώνονται με άλλες, δευτερεύουσες, υπηρεσίες για την παροχή επιπρόσθετης

ποιότητας υπηρεσιών, τη διαχείριση της εργασίας και των δεσμευμένων πόρων, και άλλες. Σε κάθε περίπτωση όλα τα παραπάνω πρέπει να προβλέπονται στις συμβάσεις μεταξύ χρήστη και παρόχου, κάτι που περιλαμβάνει επίσης τις υποχρεώσεις και τις αποζημιώσεις και από τις δύο πλευρές.

#### **4.4.1.2 Ποια αξία προστίθεται με την ενσωμάτωση της εφαρμογής σε ένα περιβάλλον πλέγματος;**

Οι φορείς παροχής υπηρεσιών που πωλούν τις υπηρεσίες μέσω των υποδομών πλέγματος προσπαθούν να κερδίσουν όσο το δυνατόν περισσότερους πελάτες υποσχόμενοι καλύτερη ποιότητα υπηρεσιών και χαμηλότερη τιμή. Γενικά αυτό είναι το κλειδί στην επιτυχία ενάντια στους ανταγωνιστές. Στην περίπτωση των εφαρμογών πολυμέσων ένας πολύ σημαντικός παράγοντας για τον πελάτη κατά την επιλογή μιας υπηρεσίας είναι η φήμη του παρόχου. Ο πελάτης επιλέγει το πάροχο υπηρεσιών που έχει διαθέσιμους πόρους, προσφέρει την απαραίτητη ποιότητα, διατηρεί τα δεδομένα ασφαλή και γενικά ακολουθεί όλους τους όρους που συμφωνήθηκαν στη σύμβαση (SLA). Επιπλέον μέσα από το πλέγμα κάποιος πάροχος υπηρεσιών θα μπορούσε να έχει κάποιο συγκριτικό πλεονέκτημα σε σχέση με τους ανταγωνιστές του μονοπωλώντας υπηρεσίες, όπως ένα ειδικό εφέ, που δεν είναι διαθέσιμες αλλού ή με άλλο τρόπο.

#### **4.4.1.3 Ποιο είναι το κόστος για την προσφορά της υπηρεσίας στον πελάτη; Πώς καθορίζονται οι τιμές;**

Σε ένα περιβάλλον πλέγματος οι περισσότερες από τις λειτουργικές απαιτήσεις των εφαρμογών πολυμέσων έχουν μεταφερθεί στους φορείς παροχής υπηρεσιών. Οι υπολογιστικοί πόροι, οι μονάδες αποθήκευσης δεδομένων, τα προγράμματα και οι άδειες χρήσης τους λειτουργούν από την πλευρά των φορέων παροχής υπηρεσιών. Συνεπώς ο φορέας παροχής υπηρεσιών στο τελικό κόστος της υπηρεσίας, που γενικά καθορίζεται μέσα από το  $\mu$  νό μους της προ φο ρίς και της ζήτησης, προσθέτει και το κόστος απόσβεσης της κάθε μονάδας κατανάλωσης που αναφέρθηκε προηγούμενα.

#### **4.4.1.4 Πώς οι πελάτες θα πληρώσουν για τις υπηρεσίες;**

Στις εφαρμογές πολυμέσων, οι πελάτες πληρώνουν για κάθε μονάδα κατανάλωσης το οποίο ανάγεται σε τρεις περιπτώσεις:

- **Πληρωμή για την χρήση της εφαρμογής.** Σε αυτό το μοντέλο ο χρήστης χρεώνεται για κάθε ολοκληρωμένη εργασία που εκτελείται στο πλέγμα. Για παράδειγμα η επεξεργασία ενός βίντεο ανάλυσης XxY με συγκεκριμένους μηχανισμούς και αλγορίθμους.
- **Πληρωμή για τη χρήση κάθε μιας από τις υπηρεσίες της εφαρμογής.** Σε αυτό το μοντέλο ο χρήστης χρεώνεται για τη χρήση των συγκεκριμένων υπηρεσιών, όπως η εξομάλυνση υφών μιας εικόνας.
- **Πληρωμή για τη χρήση των πόρων του πλέγματος.** Σε αυτό το μοντέλο ο χρήστης χρεώνεται για τη χρήση των υπολογιστικών πόρων που θα χρησιμοποιήσει.

#### **4.4.1.5 Πώς παρέχεται μια υπηρεσία;**

Οι απαιτήσεις των πελατών για την παροχή της υπηρεσίας περιλαμβάνονται στο υπογεγραμμένο SLA που έχει συμφωνηθεί. Συνεπώς οι όροι του SLA περιλαμβάνουν συγκεκριμένες υποχρεώσεις για τους παρόχους και εάν αυτοί οι όροι παραβιάζονται ο πελάτης πληρώνεται τη συμφωνηθείσα αποζημίωση. Ο στόχος του κάθε παρόχου είναι να εκτελεστεί η εργασία όπως προβλέπεται και να τηρήσει το SLA. Για αυτό το λόγο, οι πάροχοι εφαρμόζουν συγκεκριμένες πολιτικές για το πώς παρέχεται κάθε υπηρεσία προκειμένου να επιτύχουν υψηλή ποιότητα υπηρεσιών και αποτελεσματικότερη διαχείριση των πόρων.

Έχοντας λοιπόν καταλήξει στις βασικές απαιτήσεις των χρηστών και των επιχειρηματικών τους μοντέλων θα πρέπει να εξετάσουμε τις αρχιτεκτονικές και

τεχνολογικές προκλήσεις που προκύπτουν από αυτές προκειμένου να διασφαλιστεί η ποιότητα υπηρεσιών. Στο επόμενο κεφάλαιο περιγράφονται οι προκλήσεις αυτές ενώ στη συνέχεια της διατριβής αναλύουμε συγκεκριμένες αρχιτεκτονικές και τεχνολογικές λύσεις που τις αντιμετωπίζουν.

#### **4.5 Καινοτόμος υλοποίηση υπηρεσιοστρεφών υποδομών για τη διασφάλιση ποιότητας υπηρεσιών**

Στο κεφάλαιο αυτό αναλύσαμε την ανάγκη των χρηστών υπηρεσιοστρεφών υποδομών για ποιότητα υπηρεσιών και πως αυτή εκφράζεται μέσα από τα επιχειρηματικά μοντέλα. Διαπιστώσαμε πως οι ανάγκες των χρηστών και το πλαίσιο του επιχειρηματικού μοντέλου στο οποίο εντάσσεται η υποδομή είναι αυτά που καθορίζουν σε μεγάλο βαθμό τον αρχιτεκτονικό σχεδιασμό και τη λειτουργικότητά της. Στον τομέα των υπηρεσιοστρεφών υποδομών τα παραπάνω εκφράζονται μέσω των συμφωνιών επιπέδου υπηρεσιών και των μηχανισμών που τα υποστηρίζουν σε όλη τη διάρκεια ζωής της εφαρμογής μέσα στο περιβάλλον της υποδομής.

Η αποδοτική διαχείριση συμφωνιών επιπέδου υπηρεσιών είναι η πιο βασική από τις αρχιτεκτονικές αρχές σχεδιασμού που μελετήσαμε στα πλαίσια της διατριβής. Στα επόμενα κεφάλαια με βάση τα αποτελέσματα της ανάλυσης για τα επιχειρηματικά μοντέλα, την ποιότητα υπηρεσιών και τις αρχές σχεδιασμού, παρουσιάζουμε συγκεκριμένες αρχιτεκτονικές λύσεις και μηχανισμούς που διασφαλίζουν την παροχή ποιότητας υπηρεσιών σε υπηρεσιοστρεφείς υποδομές ενώ αξιολογούμε την αποτελεσματικότητά τους αναφορικά με εφαρμογές τρισδιάστατης αναπαράστασης και βιοϊατρικής.

# 5

## *Περιβάλλοντα πλέγματος επόμενης γενιάς*

Αφού παρουσιάσαμε τις αρχές λειτουργίας των πλεγμάτων στο κεφάλαιο 4, εδώ θα δούμε τη λειτουργία της εφαρμογής ψηφιακών πολυμέσων μέσα σε περιβάλλον πλέγματος επόμενης γενιάς [68] και θα παρουσιάσουμε νέες αρχιτεκτονικές προσεγγίσεις και θα τις αξιολογήσουμε μέσα από πειράματα.

Το όραμα για τα πλέγματα επόμενης γενιάς, όπως προέκυψε από το ευρωπαϊκό ερευνητικό έργο NextGRID, είναι αυτά να είναι οικονομικά βιώσιμα, και στα οποία α) τα υπάρχοντα αλλά και τυχόν νέα επιχειρηματικά μοντέλα να είναι δυνατά, β) η ανάπτυξη, η επέκταση και η συντήρηση των εφαρμογών είναι εύκολες και γ) οι επιλογές για ασφάλεια και μυστικότητα παρέχουν εμπιστοσύνη στις επιχειρήσεις, τους καταναλωτές και το ευρύτερο κοινό.

### *5.1 Εφαρμογή πολυμέσων σε περιβάλλον πλέγματος*

Η αποδοτική λειτουργία εφαρμογών πολυμέσων σε περιβάλλον πλέγματος απαιτεί για τους τελικούς χρήστες α) ευκολία στη χρήση, β) αυτοματοποιημένες διαδικασίες για την εκτέλεση ροών εργασιών και γ) διασφάλιση της παρεχόμενης ποιότητας υπηρεσιών [57][9]. Η ευκολία στη χρήση μπορεί να εξασφαλιστεί με φιλικές προς το χρήστη διεπαφές που κρύβουν όλη τη λειτουργικότητα του πλέγματος και της εφαρμογής πίσω

από καλαίσθητα GUI (Graphical User Interface). Οι διεπαφές σε κάθε περίπτωση πρέπει να ακολουθούν τις τάσεις και τεχνολογικές εξελίξεις κάθε εποχής χωρίς συμβιβασμούς αναφορικά με τη λειτουργικότητα και την παρεχόμενη ποιότητα υπηρεσιών. Μεγάλο ερευνητικό ενδιαφέρον παρουσιάζουν τα θέματα (β) και (γ) στα οποία επικεντρώνεται και η διατριβή αυτή. Οι μηχανισμοί και οι προδιαγραφές πλέγματος που παρουσιάζονται παρακάτω είναι ικανά να υποστηρίξουν την αυτοματοποίηση διαδικασίας του πλέγματος και να διασφαλίσουν υψηλό επίπεδο ποιότητας υπηρεσιών.

### **5.1.1 Χαρακτηριστικά γνωρίσματα πλέγματος επόμενης γενιάς**

Η μεγάλη πρόκληση για τα πλέγματα επόμενης γενιάς είναι να προσφέρουν στους τελικούς χρήστες ένα κατάλληλο περιβάλλον για να μπορέσουν να ικανοποιήσουν τις απαιτήσεις τους και τις απαιτήσεις των εφαρμογών τους. Δεδομένου ότι οι απαιτήσεις τόσο των τελικών χρηστών όσο και των εφαρμογών ποικίλουν, ο σκοπός δεν είναι να δημιουργήσουμε μία μοναδική αρχιτεκτονική που να μπορεί να καλύψει όλες τις ανάγκες. Αντιθέτως το ερευνητικό ενδιαφέρον επικεντρώνεται στη δημιουργία γενικών αρχιτεκτονικών αρχών (βλέπε 4.3) και μεμονωμένων πρωτοτύπων συστατικών τα οποία μπορούν να συνθέσουν πλέγματα, με δυναμικό τρόπο, ανάλογα με τις ανάγκες που υπάρχουν. Μέσω αυτών των πλεγμάτων οι τελικοί χρήστες μπορούν να έχουν πρόσβαση σε υπηρεσίες και χρησιμοποιώντας τα διαφορετικά και ευέλικτα επιχειρηματικά μοντέλα. Αυτό είναι πολύ σημαντικό για τις εφαρμογές πολυμέσων, οι οποίες έχουν επιχειρηματικό προσανατολισμό και περιλαμβάνουν διάφορους τύπους χρηστών, από τους απλούς χρήστες μέχρι τις μεγάλες επιχειρήσεις. Κοινός άξονας είναι πάντα η εξασφάλιση της ποιότητας υπηρεσιών που εκφράζεται σε όλα τα βήματα λειτουργίας της εφαρμογής μέσα από τις συμφωνίες επιπέδου υπηρεσιών.

### **5.1.2 Συμφωνίες επιπέδου υπηρεσιών**

Ο επιχειρηματικός προσανατολισμός των εφαρμογών πολυμέσων απαιτεί αυστηρά καθορισμένες σχέσεις και συμφωνίες μεταξύ των συμβαλλόμενων μερών που συμμετέχουν στη διαδικασία για χρήση υπηρεσιών. Αυτό σημαίνει ότι όλες οι αλληλεπιδράσεις μεταξύ των καταναλωτών και των φορέων παροχής υπηρεσιών, ή μεταξύ των υπηρεσιών, πρέπει να συνοδεύονται με τα αντίστοιχα SLAs. Τα SLAs έχουν πρωταρχικό ρόλο στην υποδομή και στις διαδικασίες διασφαλίζοντας αφενός τις απαιτήσεις ποιότητας υπηρεσιών για τους καταναλωτές και αφετέρου καθορίζοντας τις κατάλληλες πολιτικές χρήσης και διαχείρισης από την πλευρά των φορέων παροχής υπηρεσιών. Από την ανάλυση που πραγματοποιήθηκε για τα SLAs στις εφαρμογές πολυμέσων προέκυψαν δύο σημαντικές παρατηρήσεις.

Η πρώτη είναι ότι η γλώσσα που περιγράφει τα SLAs πρέπει να λειτουργεί σε επιχειρηματικό επίπεδο έτσι ώστε τα έγγραφα SLA και το περιεχόμενό τους να μπορούν να γίνουν κατανοητά από τους πελάτες των υπηρεσιών. Αυτός σημαίνει ότι οι όροι του SLA πρέπει να αναφέρονται σε μονάδες κατανάλωσης που συσχετίζονται με τη συγκεκριμένη εφαρμογή και με τους πόρους που απαιτούνται όπως αποθηκευτικός χώρος, μνήμη και κύκλους επεξεργασίας. Εντούτοις, οι φορείς παροχής υπηρεσιών πρέπει να γνωρίζουν και να είναι σε θέση να κάνουν την αντιστοίχιση από τους όρους υψηλού επιπέδου στις χαμηλού επιπέδου παραμέτρους υπηρεσιών. Σε αυτήν την διαδικασία, θα πρέπει να καθοριστούν και οι αντίστοιχες πολιτικές διαχείρισης (σε χαμηλό επίπεδο) προκειμένου να διασφαλιστούν οι υψηλού επιπέδου απαιτήσεις ποιότητας υπηρεσιών. Η δεύτερη παρατήρηση είναι ότι τα SLAs πρέπει να είναι αρκετά εύκαμπτα για να καλύψουν όλες τις απαιτήσεις χρηστών. Ο αριθμός και το είδος των όρων στα SLAs δεν είναι στατικός αλλά δυναμικός.

Τα SLAs που παρουσιάζουμε πιο κάτω και οι μηχανισμοί που τα υποστηρίζουν όχι μόνο καλύπτουν προαναφερθείσες απαιτήσεις αλλά μπορούν να υποστηρίξουν και συμπληρωματικές διαδικασίες χρήσης υπηρεσιών ευέλικτα και αποδοτικά. Συγκεκριμένα οι μηχανισμοί περιλαμβάνουν:

- Αποτελεσματική διαπραγμάτευση των SLA λαμβάνοντας υπόψη την αντίληψη των χρηστών αναφορικά με την ποιότητα υπηρεσίας (Quality of Experience – QoE),
- Αποτελεσματική υλοποίηση και εκτέλεση ροών εργασιών βασισμένη στις απαιτήσεις χρηστών και τα SLAs,
- Ισχυρές εγγυήσεις για την επιτυχή εκτέλεση των υπηρεσιών στο επιθυμητό επίπεδο παροχής υπηρεσιών με τη χρήση συγκεκριμένων πολιτικών διαχείρισης και λειτουργίας,
- Αποτελεσματική επίβλεψη των υπηρεσιών και της χρήσης των πόρων μέσω κοινά αποδεκτών μηχανισμών που έχουν προσυμφωνηθεί στα SLAs,
- Υποχρεώσεις και αποζημιώσεις για όλα τα εμπλεκόμενα μέρη ώστε να μπορούν να εγγυηθούν όλοι τους όρους στο έγγραφο του SLA.

### ***5.1.3 Συμφωνίες επιπέδου υπηρεσιών για εφαρμογές πολυμέσων***

Τα SLAs για τις εφαρμογές πολυμέσων περιλαμβάνουν δύο συμβαλλόμενα μέρη, τον πελάτη και τον φορέα παροχής υπηρεσιών. Ο καταναλωτής υπηρεσιών είναι η επιχείρηση παραγωγής πολυμέσων και συγκεκριμένα οι σχεδιαστές πολυμέσων. Στο σενάριο χρήσης οι επιχειρήσεις αυτές επιθυμούν πρόσβαση σε υπηρεσίες εφαρμογών μέσα από το περιβάλλον πλέγματος. Το παράδειγμα εφαρμογής πολυμέσων που εξετάσαμε είναι η τρισδιάστατη ψηφιακή αναπαράσταση (3D animation) που αποτελείται από τρεις διαδικασίες: α) τη μεταγλώττιση υφών (texture compilation), β) την



μεταγλώττιση σκιών (shaders compilation) και γ) τη τρισδιάστατη εξομάλυνση (3D rendering). Οι πρώτες δύο διαδικασίες δεν είναι υπολογιστικά «βαριές», αλλά το λογισμικό που απαιτείται και αυτές πρέπει να είναι συμβατό με το λογισμικό της εξομάλυνσης. Για τις διαδικασίες αυτές συντάξαμε τρία διαφορετικά SLAs που καθορίζουν διαφορετικά επίπεδα παροχής υπηρεσιών και κατά συνέπεια διαφορετική τιμή. Οι όροι και οι ακριβείς τιμές τους περιγράφονται στους παρακάτω πίνακες:

SLA Term	SLA Template Type		
	Bronze	Silver	Gold
Availability	0.90	0.95	0.99
Response Time (sec)	10	5	1
Price (€)	0.1	0.2	0.5
Delay Compensation per second (€) (less than response time * 2 seconds)	0.2	0.4	1.0
Failure Compensation (€)	4	8	20

*Πίνακας 1: Παράδειγμα Shader SLA*

SLA Term	SLA Template Type		
	Bronze	Silver	Gold
Availability	0.90	0.95	0.99
Response Time (sec)	10	5	2
Price (€)	0.2	0.4	0.8
Delay Compensation per second (€) (less than response time * 2 seconds)	0.4	0.8	1.6
Failure Compensation (€)	8	16	40

*Πίνακας 2: Παράδειγμα Texture SLA*

Σε αυτά τα παραδείγματα ο χρήστης μπορεί να επιλέξει τον κατάλληλο τύπο SLA που ικανοποιεί τις ανάγκες του. Οι όροι παροχής ποιότητας υπηρεσιών για αυτές τις υπηρεσίες είναι η διαθεσιμότητα (availability) και ο χρόνος απόκρισης (response time). Συγκεκριμένα οι χρήστες απαιτούν υψηλή διαθεσιμότητα για την υπηρεσία και χαμηλό χρόνο απόκρισης. Η διαθεσιμότητα απεικονίζει το «uptime» της υπηρεσίας που

προσφέρεται μέσω του SLA. Για παράδειγμα η διαθεσιμότητα με τιμή 0.90 στο «χάλκινο» SLA δηλώνει ότι εάν ένας πελάτης προσπαθεί να δεσμεύσει μια υπηρεσία, η πιθανότητα της επιτυχίας είναι 90%. Αυτός ο όρος δεν συνδέεται άμεσα με κάποιο όρο αποζημίωσης αλλά είναι ένα χαρακτηριστικό που χρησιμοποιείται για να επιλεγθεί το καλύτερο SLA για τον πελάτη. Να σημειώσουμε επίσης ότι η διαθεσιμότητα δεν μπορεί να μετρηθεί από τον πελάτη αλλά είναι ένα χαρακτηριστικό για το οποίο ο φορέας παροχής υπηρεσιών πιστοποιείται από έναν τρίτο. Ο χρόνος απόκρισης μετριέται από το φορέα παροχής υπηρεσιών. Εντούτοις, ο τελικός χρήστης μπορεί να αντιληφτεί τυχών ανακρίβειες με την καθυστέρηση στη λήψη του αποτελέσματος. Κάθε τύπος SLA έχει φυσικά διαφορετικό κόστος και αποζημιώσεις οι οποίες βασίζονται στην ποιότητα υπηρεσίας που προσφέρεται. Υπάρχουν δύο τύποι αποζημιώσεων, ένας για την καθυστέρηση στη λήψη του αποτελέσματος και ένας για την αποτυχία εργασίας. Στο παράδειγμα αυτό θεωρούμε πως εάν το αποτέλεσμα καθυστερήσει περισσότερες από δύο φορές του χρόνου απόκρισης, η εργασία θεωρείται ως αποτυχημένη και ο φορέας παροχής υπηρεσιών πρέπει να πληρώσει στον πελάτη τη μέγιστη αποζημίωση καθυστέρησης συν την αποζημίωση αποτυχίας.

Το SLA για τη διαδικασία του rendering είναι αρκετά πιο περίπλοκα. Οι μονάδες στη διαδικασία αυτή είναι οι εικόνες-πλαίσια (frames) που εξομαλύνονται. Αυτό αναδεικνύει το πρόβλημα που θίξαμε προηγούμενα σχετικά με την αντιστοίχιση μεταξύ των απαιτήσεων υψηλού επιπέδου και των μονάδων κατανάλωσης και χαμηλού επιπέδου παραμέτρων από την πλευρά του παρόχου. Να σημειώσουμε ότι χρόνος επεξεργασίας για κάθε πλαίσιο ποικίλει από λιγότερο από ένα δευτερόλεπτο σε αρκετές ώρες ανάλογα με την πολυπλοκότητα της εργασίας. Αυτό έχει επιπτώσεις σε όλες τις άλλες παραμέτρους ποιότητας όπως ο χρόνος απόκρισης και κάνει την αντιστοίχιση των παραμέτρων μια πολύ δύσκολη και σύνθετη διαδικασία.

Μετά από μελέτη των εφαρμογών πολυμέσων σε περιβάλλον πλέγματος αποφασίσαμε την ταξινόμηση των frames ανάλογα με την πολυπλοκότητά τους και κατά συνέπεια ανάλογα με τις υπολογιστικές απαιτήσεις τους. Για να είναι η κατηγοριοποίηση όσο το δυνατόν γενικότερη πραγματοποιήσαμε ένα πείραμα. Ένας υπολογιστής με τα ακόλουθα χαρακτηριστικά χρησιμοποιήθηκε ως πρωτότυπο:

<b>CPU</b>	Intel P4 3GHz
<b>RAM Memory</b>	1GB
<b>Renderer</b>	SitexGraphics AIR 5.0

*Πίνακας 3: Προδιαγραφές πρότυπης μηχανής*

Οι επιδόσεις της συγκεκριμένης μηχανής έχουν αξιολογηθεί με συγκεκριμένες μετρήσεις ενώ κάθε άλλη μηχανή μπορεί να ανάγει τις επιδόσεις της σε αυτή με τους κατάλληλους πολλαπλασιαστές. Για το πείραμα, διάφορες εργασίες τρισδιάστατης απεικόνισης κάθε μία από τις οποίες αποτελείται από 300 πλαίσια εκτελέστηκαν χρησιμοποιώντας αυτήν την μηχανή. Με βάση το χρόνο επεξεργασίας τους προέκυψαν 10 κατηγορίες πολυπλοκότητας. Αυτή η κατηγοριοποίηση χρησιμοποιήθηκε στα SLAs για την αντιστοίχιση των απαιτήσεων υψηλού επιπέδου των πελατών στις χαμηλού επιπέδου παραμέτρους των φορέων παροχής υπηρεσιών:

<b>Category</b>	<b>Processing Time (seconds)</b>	
	<b>Minimum</b>	<b>Maximum</b>
1		1
2	1	3
3	3	6
4	6	10
5	10	18
6	18	30
7	30	45
8	45	70
9	70	100
10	100	

*Πίνακας 4: Κατηγορίες πολυπλοκότητας*

Έτσι οι πελάτες χρησιμοποιώντας την κατηγοριοποίηση αυτή, εντάσσουν την εργασία τους σε μια κατηγορία πολυπλοκότητας ώστε να επιλεγεί μετέπειτα το κατάλληλο SLA. Τα SLAs όπως και στις προηγούμενες υπηρεσίες περιλαμβάνουν δύο τύπους όρων, τους στατικούς και τους μεταβλητούς. Οι στατικοί όροι καθορίζουν τον τύπο υπηρεσιών ενώ οι μεταβλητοί την απόδοσή τους. Οι στατικοί όροι για αυτή την υπηρεσία είναι ο τύπος renderer, η έκδοση του, η ανάλυση εικόνας και εάν το SLA μπορεί να δεχτεί μια συνολική αποτυχία εργασίας. Ο τελευταίος όρος σημαίνει ότι εάν ένα πλαίσιο αποτύχει, ο φορέας παροχής υπηρεσιών πρέπει να πληρώσει αποζημίωση για ολόκληρη την εργασία. Οι μεταβλητοί όροι αναφέρονται στην κατηγορία πολυπλοκότητας, τη διαθεσιμότητα, το χρόνο απόκρισης, την αξία και τις αποζημιώσεις σε περίπτωση αποτυχίας. Στον ακόλουθο πίνακα παρουσιάζονται παραδείγματα για αυτό το SLA:

SLA Term		SLA Template Type		
		Bronze	Silver	Gold
<b>Renderer Type</b>		Sitexgraphics AIR		
<b>Renderer Version</b>		5.0		
<b>Resolution</b>		640x480		
<b>Availability</b>		0,90	0,95	0,99
<b>Accepts Total Failure</b>		no	no	yes
<b>Cat1</b>	Response Time per Frame (sec)	5	2	1
	Price per frame (€)	0,004	0,01	0,02
	Delay Compensation per Frame per second (less than response time * 2 seconds) (€)	0,008	0,02	0,1
	Failure Compensation (Delay more than response time * 2) (€)	0,08	0,2	1
<b>Cat2</b>	Response Time per Frame (sec)	15	6	3
	Price per frame (€)	0,012	0,03	0,06
	Delay Compensation per Frame per second (less than response time * 2 seconds) (€)	0,024	0,06	0,3
	Failure Compensation (Delay more than response time * 2) (€)	0,24	0,6	3
<b>Cat3</b>	Response Time per Frame (sec)	30	12	6
	Price per frame (€)	0,024	0,06	0,12
	Delay Compensation per Frame per second (less than response time * 2 seconds) (€)	0,048	0,12	0,6
	Failure Compensation (Delay more than response time * 2) (€)	0,48	1,2	6
<b>Cat4</b>	Response Time per Frame (sec)	50	20	10
	Price per frame (€)	0,04	0,1	0,2

	Delay Compensation per Frame per second (less than response time * 2 seconds) (€)	0,08	0,2	1
	Failure Compensation (Delay more than response time * 2) (€)	0,8	2	10
<b>Cat5</b>	Response Time per Frame (sec)	90	36	18
	Price per frame (€)	0,072	0,18	0,36
	Delay Compensation per Frame per second (less than response time * 2 seconds) (€)	0,144	0,36	1,8
	Failure Compensation (Delay more than response time * 2) (€)	1,44	3,6	18
<b>Cat6</b>	Response Time per Frame (sec)	150	60	30
	Price per frame (€)	0,12	0,3	0,6
	Delay Compensation per Frame per second (less than response time * 2 seconds) (€)	0,24	0,6	3
	Failure Compensation (Delay more than response time * 2) (€)	2,4	6	30
<b>Cat7</b>	Response Time per Frame (sec)	225	90	45
	Price per frame (€)	0,18	0,45	0,9
	Delay Compensation per Frame per second (less than response time * 2 seconds) (€)	0,36	0,9	4,5
	Failure Compensation (Delay more than response time * 2) (€)	3,6	9	45
<b>Cat8</b>	Response Time per Frame (sec)	350	140	70
	Price per frame (€)	0,28	0,7	1,4
	Delay Compensation per Frame per second (less than response time * 2 seconds) (€)	0,56	1,4	7
	Failure Compensation (Delay more than response time * 2) (€)	5,6	14	70
<b>Cat9</b>	Response Time per Frame (sec)	500	200	100
	Price per frame (€)	0,4	1	2
	Delay Compensation per Frame per second (less than response time * 2 seconds) (€)	0,8	2	10
	Failure Compensation (Delay more than response time * 2) (€)	8	20	100
<b>Cat10</b>	Response Time per Frame (sec)	2500	1000	500
	Price per frame (€)	2	5	10
	Delay Compensation per Frame per second (less than response time * 2 seconds) (€)	4	10	50
	Failure Compensation (Delay more than response time * 2) (€)	40	100	500

**Πίνακας 5: Παράδειγμα Rendering SLA**

Αυτοί οι όροι και οι τιμές τους συντάσσονται σε κατάλληλα έγγραφα xml ακολουθώντας συγκεκριμένα πρότυπα, ενώ τα δομικά στοιχεία του πλέγματος είναι σε θέση να τα

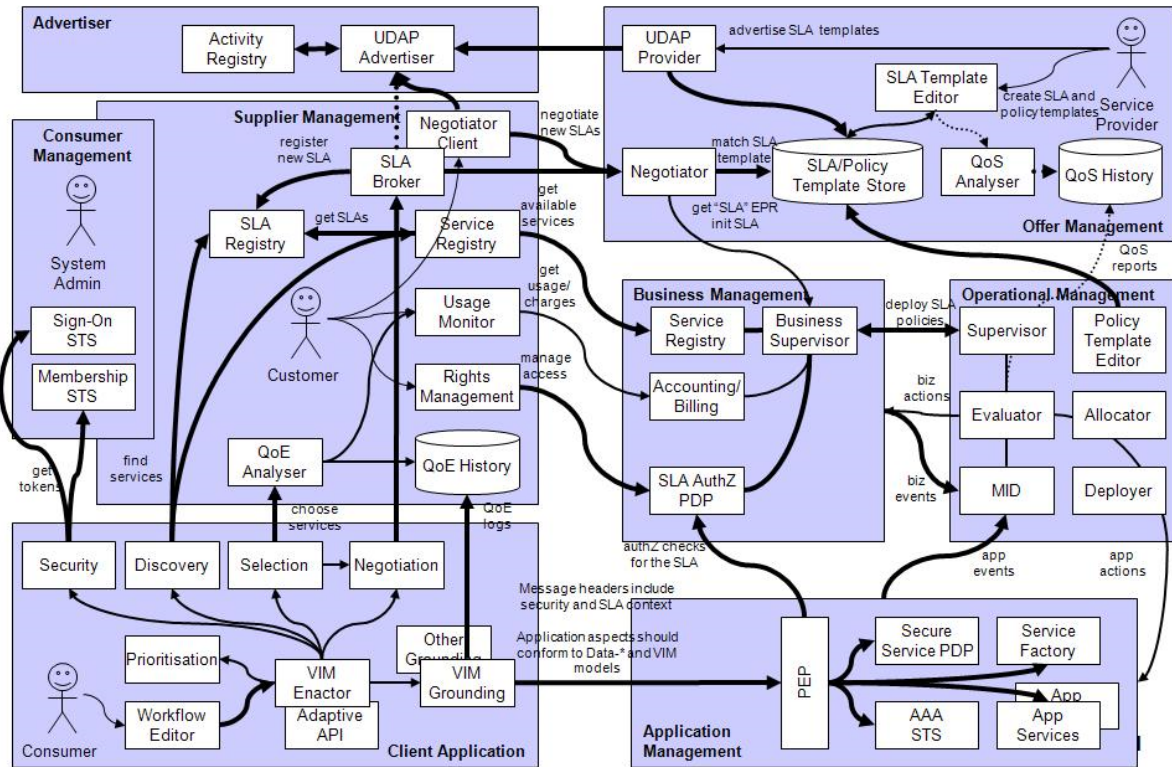
μεταφράσουν και να πάρουν τις αποφάσεις που απαιτούνται κάθε φορά μέσα από αυτοματοποιημένες διαδικασίες.

#### **5.1.4 Ο ρόλος των ροών εργασιών**

Οι εφαρμογές που εντάσσονται και χρησιμοποιούν πλέγματα και υπηρεσιοστρεφείς υποδομές τις περισσότερες φορές δεν είναι απαιτούν μόνο μια υπηρεσία για τη ολοκληρωμένη λειτουργία τους αλλά μία σειρά από υπηρεσίες που αλληλεπιδρούν μεταξύ τους και εκτελούνται είτε παράλληλα είτε σειριακά [24]. Οι εφαρμογές όπως και γενικά οι κάθε λογής απαιτήσεις των χρηστών εκφράζονται μέσα από ροές εργασιών (workflows) τα οποία περιγράφουν σε μια γλώσσα που μπορεί να γίνει κατανοητή από την υποδομή την εργασία που θέλουν να εκτελέσουν. Αυτό απαιτεί από την πλευρά της υποδομής μηχανισμούς με τις ικανότητες αναζήτησης υπηρεσιών, επιλογής τους και τελικά εκτέλεσής τους. Οι ροές εργασίας είναι στενά συνδεδεμένες με το επιχειρηματικό μοντέλο της εφαρμογής και τις απαιτήσεις χρηστών για ποιότητα υπηρεσίας. Λειτουργούν λοιπόν συνδυαστικά με τα SLA, και τις πολιτικές λειτουργίας και διαχείρισης των εργασιών.

## **5.2 Γενικό διάγραμμα ενσωμάτωσης**

Με βάση τις αρχές σχεδιασμού που αναλύσαμε σε προηγούμενο κεφάλαιο, τις επιχειρηματικές απαιτήσεις για ποιότητα υπηρεσιών καταλήξαμε σε γενική αρχιτεκτονική για τα περιβάλλοντα πλέγματος που περιλαμβάνει όλα τα συστατικά στοιχεία (components) καθώς και τις αλληλεπιδράσεις μεταξύ τους. Η αρχιτεκτονική αυτή συνοψίζεται στο παρακάτω διάγραμμα ενσωμάτωσης (integration diagram) που αναπτύξαμε στα πλαίσια του ευρωπαϊκού ερευνητικού έργου NextGRID [68][67][66].



Σχήμα 15: Διάγραμμα ενσωμάτωσης

Τα συνθετικά στοιχεία του πλέγματος έχουν κατηγοριοποιηθεί ανάλογα με το ρόλο τους στην υποδομή και το τομέα στον οποίο ανήκουν (πάροχο ή καταναλωτή) ενώ με (έντονα) βέλη απεικονίζονται οι (βασικές) αλληλεπιδράσεις τους. Να σημειώσουμε ότι η παραπάνω εικόνα αποτελεί τη γενικότερη αρχιτεκτονικά περιγραφή της υποδομής και δεν απαιτούνται όλα αυτά τα συστατικά ώστε το πλέγμα να είναι λειτουργικό και να ικανοποιήσει τις λειτουργικές ανάγκες της υποδομής.

### 5.3 Αρχιτεκτονική πλέγματος για εφαρμογές πολυμέσων

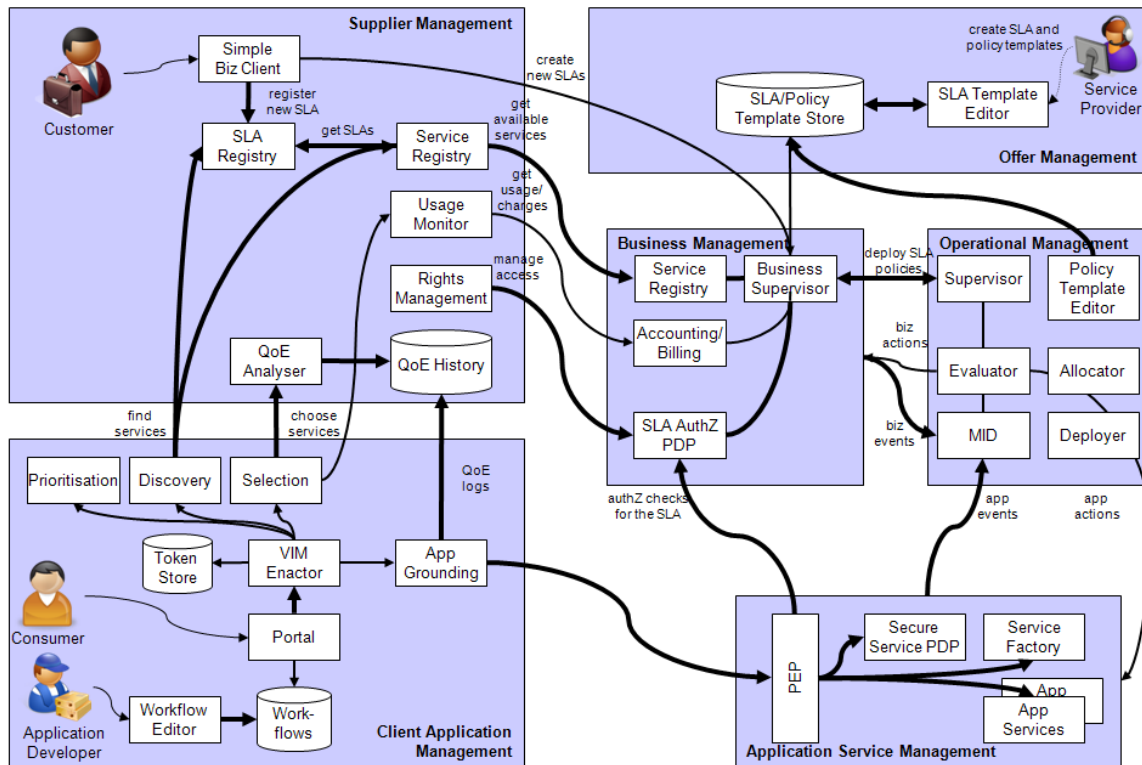
#### 5.3.1 Ενσωμάτωση στοιχείων πλέγματος

Οι εφαρμογές πολυμέσων είναι λόγω της επιχειρηματικής φύσης τους ιδανικά για την αξιολόγηση της γενικής αρχιτεκτονικής πλέγματος και των προδιαγραφών που παρουσιάστηκαν προηγουμένα μέσα από πειράματα. Μετά από ανάλυση των απαιτήσεων των εφαρμογών πολυμέσων καταλήξαμε στα απαραίτητα συστατικά για την ικανοποίηση

*Η σελίδα αυτή είναι σκόπιμα λευκή*



των απαιτήσεών τους. Το ακόλουθο διάγραμμα συστατικών απεικονίζει τη γενική αρχιτεκτονική πλέγματος για εφαρμογές πολυμέσων, τα συστατικά που χρησιμοποιούνται και τις αλληλεπιδράσεις τους:



Σχήμα 16: Διάγραμμα ενσωμάτωσης για εφαρμογές πολυμέσων

Τα βασικά δομικά στοιχεία της αρχιτεκτονικής είναι:

- Ο διαχειριστής ροών εργασιών (VIM Enactor)
- Ο μηχανισμός επιλογής υπηρεσιών (Selection)
- Το μητρώο καταγραφής ποιότητας υπηρεσιών (QoE History)
- Η μηχανισμός ανάλυσης ποιότητας υπηρεσιών (QoE Analyser)
- Το μητρώο καταγραφής υπηρεσιών (Service Registry)
- Το μητρώο καταγραφής συμφωνιών επιπέδου υπηρεσιών (SLA Registry)

*Η σελίδα αυτή είναι σκόπιμα λευκή*

- Ο μηχανισμός επιχειρησιακής διαχείρισης (Business Management)
- Ο μηχανισμός λειτουργικής διαχείρισης (Operational Management)

Η διεπαφή της εφαρμογής και κατά συνέπεια του πλέγματος για τους χρήστες είναι μια διαδικτυακή πύλη η οποία επικοινωνεί με τις υπηρεσίες πλέγματος μέσω του VIM Enactor. Μέσω της πύλης ο χρήστης είναι σε θέση να επιλέξει την κατάλληλη υπηρεσία ή ροή εργασίας, να καθορίσει τις απαιτήσεις του, όπως την τιμή και τον επιθυμητό χρόνο εκτέλεσης, και να τις υποβάλει στο πλέγμα για αυτοματοποιημένη εκτέλεση.

Από την πλευρά του παρόχου χρησιμοποιήθηκε μια προσαρμοσμένη έκδοση των βασικών υπηρεσιών GRIA [43] για την εκτέλεση των εργασιών που υποστηρίζει την επικοινωνία μεταξύ του περιβάλλοντος εκτέλεσης και των τμημάτων λειτουργικής και επιχειρησιακής διαχείρισης. Η πιο σημαντική διαφοροποίηση της συγκεκριμένης έκδοσης είναι ότι προσαρμόστηκε ώστε να μπορεί να υποστηρίξει τα SLAs που περιγράφηκαν προηγουμένως και να είναι σε θέση να διαχειριστεί τις εργασίες και τους πόρους με βάση αυτά. Οι πολιτικές διαχείρισης είναι βασισμένες στις απαιτήσεις SLA ενώ τα τμήματα κώδικα που χρησιμοποιούνται για την εκτέλεση των εφαρμογών (wrappers) υποβάλλουν συνεχώς αναφορές για την κατάσταση εκτέλεσης. Το ακόλουθο xml είναι ένα παράδειγμα ενημέρωσης από το σύστημα εκτέλεσης της εργασίας προς το μηχανισμό λειτουργικής διαχείρισης:

```
<?xml version="1.0" encoding="UTF-8"?>
<Event ReportTime="2001-12-31T12:00:00"
  xmlns:muws-1="http://docs.oasis-open.org/wsdm/2004/12/muws/wsdm-muws-
part1.xsd"
  xmlns:ng-e="http://www.nextgrid.org/2006/08/event"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <ng-e:EventId>
    http://anEventId
  </ng-e:EventId>
  <ng-e:SourceComponent>
    <muws-1:ResourceId>
      http://aResourceId
    </muws-1:ResourceId>
    <muws-1:ComponentAddress>
      <epr>http://aService</epr>
```

```

    </muws-1:ComponentAddress>
</ng-e:SourceComponent>
<ng-e:ReporterComponent>
  <muws-1:ResourceId>
    http://aResourceId
  </muws-1:ResourceId>
  <muws-1:ComponentAddress>
    <epr>http://aService</epr>
  </muws-1:ComponentAddress>
</ng-e:ReporterComponent>
<ng-e:EventGroupID>
  http://www.gria.org/sla/metric/activity/current-activities
</ng-e:EventGroupID>
<ng-e:SequenceNumber>
  0
</ng-e:SequenceNumber>
<ns10:current-activities
  xmlns:ns10="http://it-innovation.soton.ac.uk/2005/grid">
  1.0
</ns10:current-activities>
<ns11:BillingInformation
  xmlns:ns11="http://it-innovation.soton.ac.uk/2005/grid">
  <ns12:Address
    xmlns:ns12="http://www.w3.org/2005/08/addressing">
    https://mensang.it-
innovation.soton.ac.uk:8443/bizmgt/services/SlaResource
  </ns12:Address>
  <ns13:ReferenceParameters
    xmlns:ns13="http://www.w3.org/2005/08/addressing">
    <itinnov:ConversationID
      wsu:Id="id-31790249"
      xmlns:itinnov="http://it-innovation.soton.ac.uk/2005/grid"
      xmlns:wsu="http://docs.oasis-open.org/wss/2004/01/oasis-200401-
wss-wssecurity-utility-1.0.xsd">
      40894e1f-14acf988-0114-acf9a0bd-0001
    </itinnov:ConversationID>
  </ns13:ReferenceParameters>
  <ns14:Metadata
    xmlns:ns14="http://www.w3.org/2005/08/addressing">
    <ns4:title
      xmlns:ns4="http://purl.org/dc/elements/1.1/">
      Sla
      (40894e1f-14acf988-0114-acf9a0bd-0001)
    </ns4:title>
    <ns5:serviceresourcetype
      xmlns:ns5="http://it-innovation.soton.ac.uk/2005/grid">
      http://www.it-innovation.soton.ac.uk/nextgrid/resource/sla-
resource
    </ns5:serviceresourcetype>
    <ns7:sla-template-type
      xmlns:ns7="http://www.nextgrid.org">
      a-slat-type
    </ns7:sla-template-type>
    <ns1:type
      xmlns:ns1="http://it-innovation.soton.ac.uk/2005/grid">
      uk.ac.soton.itinnovation.nextgrid.comms.bizmgt.SlaConversation
    </ns1:type>
    <ns1:ManagementType
      xmlns:ns1="http://it-innovation.soton.ac.uk/2006/grid">
      Business Management Service
    </ns1:ManagementType>

```

```
</ns14:Metadata>  
</ns11:BillingInformation>  
</Event>
```

**Σχήμα 17: Παράδειγμα μηνύματος ενημέρωσης**

Επιπλέον, συμπληρωματικά στην υποδομή χρησιμοποιήθηκαν τα ακόλουθα συστατικά:

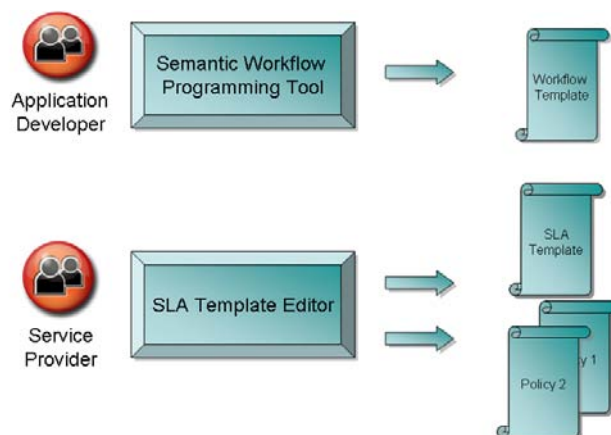
- Εργαλείο προγραμματισμού ροών εργασιών και
- Συντάκτης SLA και πολιτικών διαχείρισης.

Αυτά τα δύο συστατικά παρέχουν μέσω ενός φιλικού και απλού GUI την απαραίτητη λειτουργικότητα για τη δημιουργία, συμβατών με τα πρότυπα, ροών εργασιών, SLAs και πολιτικών διαχείρισης.

### 5.3.2 Ακολουθία λειτουργίας

Οι κύριοι δράστες στα σενάρια λειτουργίας των εφαρμογών πολυμέσων είναι οι παρακάτω:

- Υπεύθυνος για την ανάπτυξη εφαρμογής
- Πάροχος υπηρεσιών
- Πελάτης
- Καταναλωτής

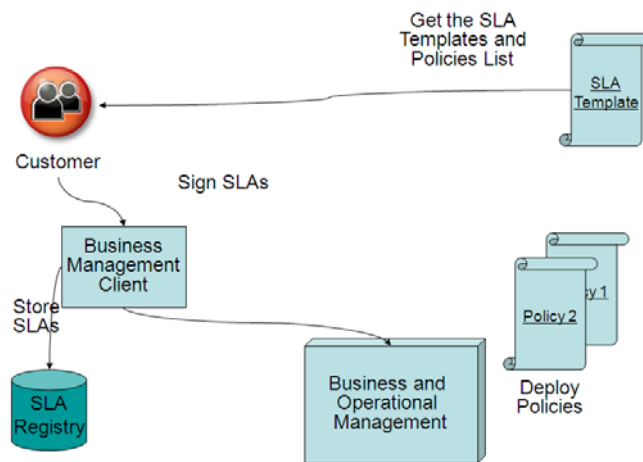


**Σχήμα 18: Δημιουργία προτύπων ροών εργασιών, συμφωνιών επιπέδου υπηρεσιών και πολιτικών διαχείρισης**

*Η σελίδα αυτή είναι σκόπιμα λευκή*

Αρχικά, χρησιμοποιείται το εργαλείο προγραμματισμού ροών εργασιών και ο συντάκτης SLAs / πολιτικών διαχείρισης από την υπεύθυνο ανάπτυξης εφαρμογής για την δημιουργία των απαραίτητων προτύπων για την εφαρμογή (Σχήμα 18).

Ο πελάτης της εφαρμογής πολυμέσων συμφωνεί σχετικά με τα SLAs για την παροχή των υπηρεσιών και έπειτα «φορτώνει» τα διαθέσιμα πρότυπα ροών εργασιών στην πύλη υποβολής εργασίας. (Σχήμα 19)

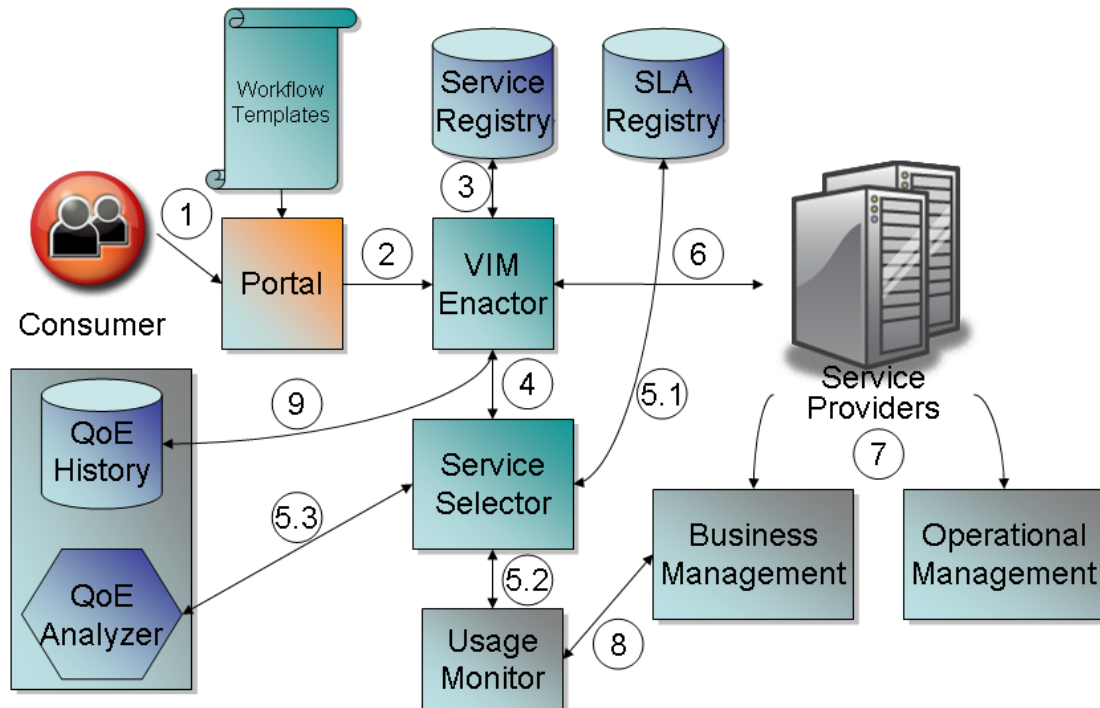


Σχήμα 19: Λειτουργίες καταναλωτή

Ο κύριος χρήστης των υπηρεσιών NextGRID είναι ο καταναλωτής (Σχήμα 20). Ο καταναλωτής μέσα από την πύλη ξεκινά μια ροή εργασίας για την εφαρμογή του και καθορίζει όλες τις απαιτήσεις για την εκτέλεση της εργασίας, όπως για παράδειγμα το χρόνος απόκρισης. (1)

*Η σελίδα αυτή είναι σκόπιμα λευκή*





Σχήμα 20: Αρχιτεκτονική και αλληλουχία χρήσης

Στη συνέχεια, η πύλη επικοινωνεί με το VIM Enactor (2) και πραγματοποιείται η αναζήτηση των κατάλληλων υπηρεσιών που ικανοποιούν τη ροή εργασίας και τους περιορισμούς που έχουν τεθεί (3). Μέσω του μηχανισμού επιλογής υπηρεσιών (4) οι καλύτερες διαθέσιμες υπηρεσίες επιλέγονται (5). Συγκεκριμένα, ο μηχανισμός επιλογής υπηρεσιών επικοινωνεί με το μητρώο SLAs για να βρει διαθέσιμα SLAs για κάθε υπηρεσία (5.1), ελέγχει εάν υπάρχουν διαθέσιμοι πόροι (5.2) και τελικά επικοινωνεί με το μηχανισμό ανάλυσης ποιότητας υπηρεσιών για να δει τις επιδόσεις των υπηρεσιών στο παρελθόν αναφορικά με την ποιότητα υπηρεσίας τους (5.3). Η διαδικασία συνεχίζεται με την εκτέλεση της ροής εργασίας στην υποδομή των επιλεγμένων φορέων παροχής υπηρεσιών (6). Η εκτέλεση εργασίας ελέγχεται συνεχώς και οι κατάλληλες πολιτικές διαχείρισης εφαρμόζονται όταν προκύψει κάποιο απρόβλεπτο ή όχι παραστατικό (7). Τέλος η χρήση της υπηρεσίας και του SLA καταγράφονται (8) όπως επίσης και η ποιότητα των υπηρεσιών που προσφέρθηκε σε κάθε εκτέλεση (9).

*Η σελίδα αυτή είναι σκόπιμα λευκή*

## **5.4 Πειράματα αξιολόγησης**

Για την αξιολόγηση των επιδόσεων της προτεινόμενης αρχιτεκτονικής και των μηχανισμών διασφάλισης των μηχανισμών ποιότητας υπηρεσιών σε περιβάλλον πλέγματος επόμενης γενιάς πραγματοποιήθηκαν δύο πειράματα:

- Αξιολόγηση επιπέδου ποιότητας υπηρεσιών στην ψηφιακή εφαρμογή μέσων
- Δρομολόγηση ροών εργασιών με βάση τις συμφωνίες επιπέδου υπηρεσιών

### **5.4.1 Πείραμα 1: Αξιολόγηση επιπέδου ποιότητας υπηρεσιών στην ψηφιακή εφαρμογή μέσων**

Στο πείραμα αυτό διερευνήσαμε την λειτουργία των μηχανισμών αξιολόγησης του επιπέδου παροχής υπηρεσιών. Οι πληροφορίες που προσφέρουν οι μηχανισμοί αυτή υποστηρίζουν την αποδοτική επιλογή φορέων παροχής υπηρεσιών για εφαρμογές πολυμέσων. Οι εφαρμογές πολυμέσων είναι αρκετά απαιτητικές από άποψη της ποιότητας υπηρεσιών και η διαχείρισή της ένα ιδιαίτερα δύσκολο και απαιτητικό έργο εφόσον είναι το βασικό κριτήριο για την επιλογή ή όχι μιας υπηρεσίας. Σε προηγούμενες υποδομές ο πελάτης έπρεπε να βασίσει την απόφασή του απλώς στην εμπειρία και στο τι οι πάροχοι «διαφημίζουν». Στην υποδομή που παρουσιάσαμε η εμπειρία αυτή πολιτικοποιείται και καταγράφεται υποστηρίζοντας τον πελάτη στην επιλογή της υπηρεσίας. Στα πλαίσια του πειράματος ερευνήσαμε την αποτελεσματικότητα και τη δυνατότητα εφαρμογής των μηχανισμών αξιολόγησης ποιότητας υπηρεσίας σε ένα πραγματικό σενάριο.

### **5.4.2 Αποτελέσματα**

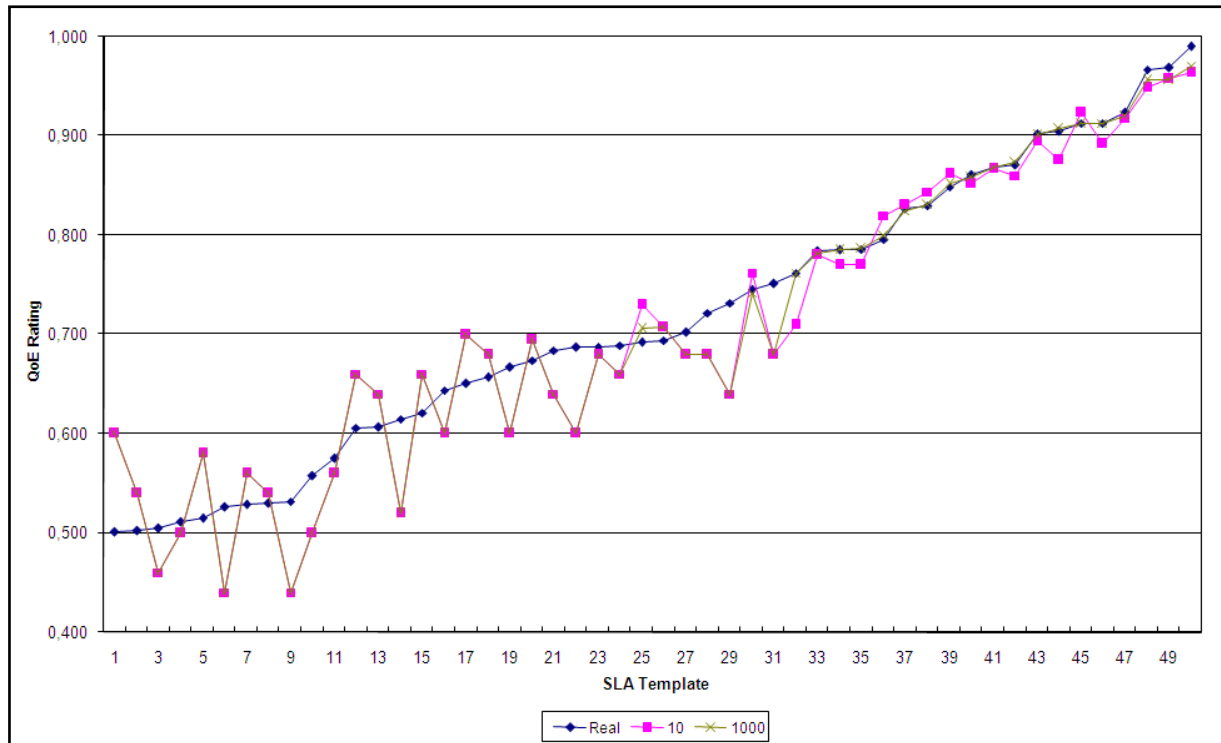
Μέσα στο πειραματικό περιβάλλον πλέγματος για εφαρμογές πολυμέσων θεωρήσαμε διάφορους παρόχους που προσφέρουν υπηρεσίες. Κάθε πάροχος υπηρεσιών μπορούσε να προσφέρει τις υπηρεσίες τους κάτω από περισσότερους του ενός τύπους SLA. Αυτά τα SLA διαφοροποιούνταν στους όρους που αφορούσαν την ποιότητα υπηρεσίας και φυσικά

στην τιμή. Επιπρόσθετα κάθε ένα από αυτά είχε διαφορετική πιθανότητα αποτυχίας, κάτι που ήταν προφανώς ήταν άγνωστο στο χρήστη.

Αρχικά το σύστημα λειτούργησε με τυχαία επιλογή υπηρεσιών που μπορούσαν να ικανοποιήσουν τις απαιτήσεις του χρήστη χωρίς να λαμβάνονται υπόψη οι συστάσεις από τους μηχανισμούς αξιολόγησης ποιότητας υπηρεσιών. Στη συνέχεια η επιλογή υπηρεσιών πραγματοποιήθηκε με βάση τις συστάσεις από τους μηχανισμούς αξιολόγησης. Στην περίπτωση αυτή, το SLA και η υπηρεσία που επιλέγεται είχε παρουσιάσει καλύτερη συμπεριφορά στο παρελθόν και έτσι είναι πιθανότερο να παρουσιάσει το ίδιο καλή συμπεριφορά και για τη συγκεκριμένη εργασία σε σχέση με τα ανταγωνιστικά. Για το πείραμα χρησιμοποιήσαμε 50 τύπους SLAs κάθε ένα από τα οποία είχε μέση πιθανότητα επιτυχίας από 50% έως 100%. Επιπλέον δεδομένου ότι η αποτυχία σχετίζεται και με την ίδια τη εργασία θεωρήσαμε ότι η παραπάνω πιθανότητα μπορεί σε μια εργασία να μεταβάλλεται  $\pm 10\%$ .

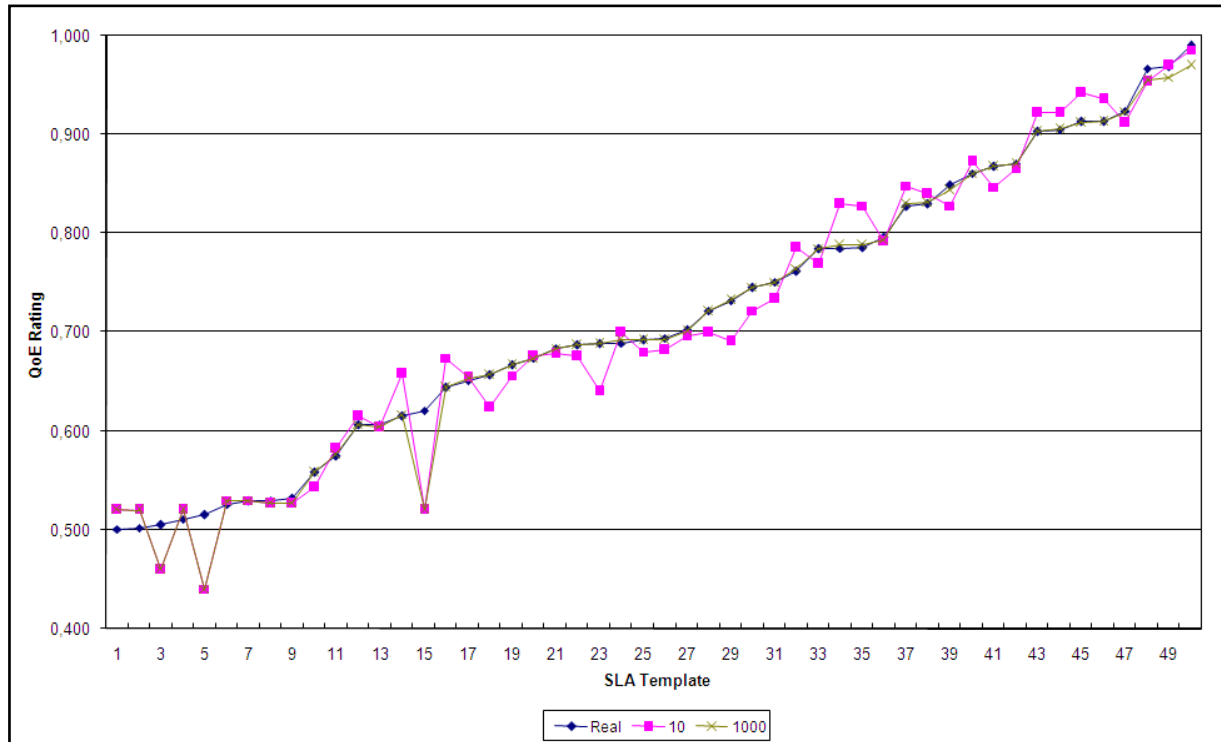
Για τις ανάγκες του πειράματος οι wrappers εκτέλεσης των εργασιών προσαρμόστηκαν προκειμένου να δημιουργούν τεχνητές αποτυχίες και κατά συνέπεια παραβιάσεις για τα SLAs. Το πείραμα πραγματοποιήθηκε με πέντε διαφορετικές ομάδες εργασιών προς εκτέλεση. Κάθε ομάδα είχε διαφορετικό αριθμό τρισδιάστατων πλαισίων απεικόνισης που μεταφράζεται σε διαφορετικό όγκο δουλειάς και διαφορετική ανάγκη σε υπηρεσίες πλέγματος που ισοδυναμούν με αιτήσεις για SLAs. Επομένως μεγαλύτερος όγκος δουλειάς σημαίνει περισσότερα SLA με παρόχους για να εκτελέσουν την ομάδα εργασίας. Συγκεκριμένα, οι ομάδες εργασίας απαιτούν 10%, 20%, 40%, 80% και 100% των διαθέσιμων υπηρεσιών και κατά συνέπεια των SLAs. Με αυτόν τον τρόπο, οι μηχανισμοί αξιολόγησης της ποιότητας υπηρεσιών εξετάζονται κάτω από διαφορετικό φόρτο εργασίας για τους παρόχους.

Τα ακόλουθα διαγράμματα (Σχήμα 21, Σχήμα 22) παρουσιάζουν την εκτίμηση ποιότητας υπηρεσιών (από 0 έως 1) για κάθε SLA. Στο πρώτο διάγραμμα ο φόρτος της υποδομής είναι 40% ενώ στο δεύτερο είναι 80%.



Σχήμα 21: Αξιολόγηση για χρήση 40%

*Η σελίδα αυτή είναι σκόπιμα λευκή*



Σχήμα 22: Αξιολόγηση για χρήση 80%

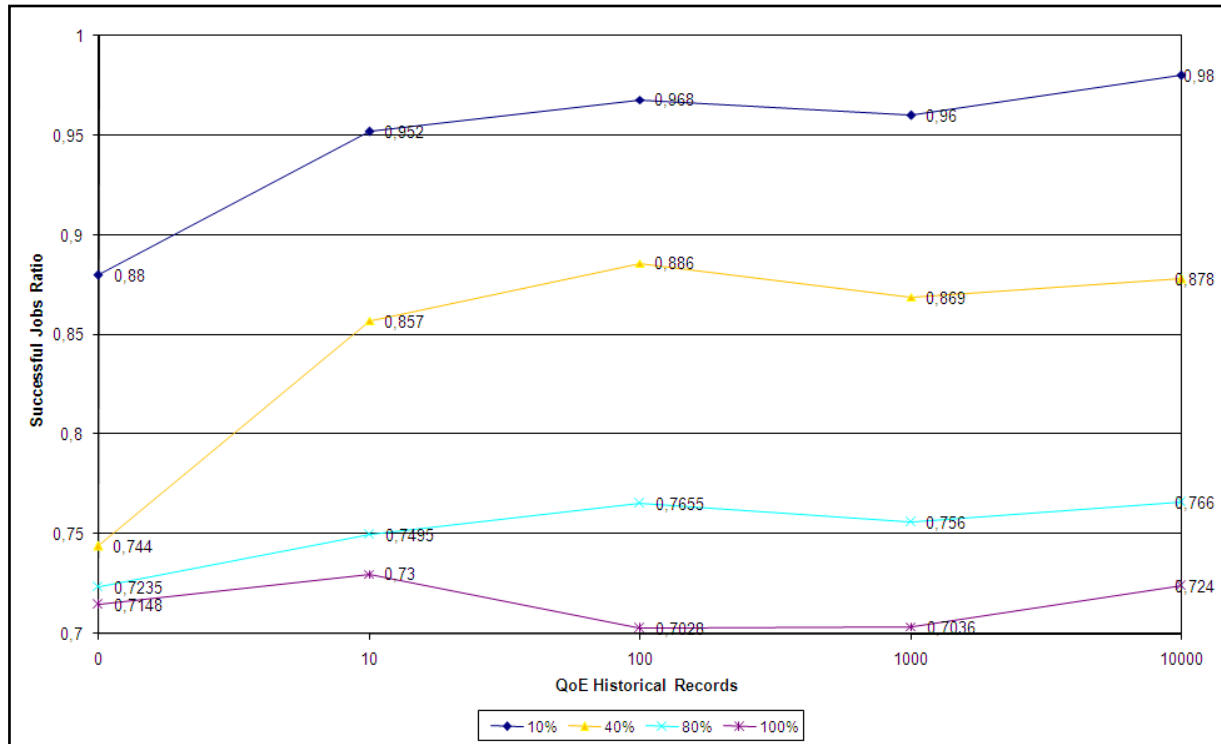
Η μπλε γραμμή δείχνει την πραγματική τιμή για την πιθανότητα αποτυχίας για κάθε SLA ενώ η κόκκινη και η κίτρινη, την τιμή που πρόεκυψε από τους μηχανισμούς αξιολόγησης ποιότητας υπηρεσίας μετά από 10 και μετά από 1000 εκτελέσεις εργασιών. Να σημειώσουμε ότι η κάθε επιλογή εκτέλεσης βασίζεται στις συστάσεις από τους μηχανισμούς αξιολόγησης και στα αποτελέσματα των προηγούμενων εργασιών που εκτελέστηκαν. Παραδείγματος χάριν, το SLA που χρησιμοποιήθηκε για την εκτέλεση της εργασίας 1000 βασίζεται στην αξιολόγηση των προηγούμενων 999. Μπορούμε να δούμε από τα διαγράμματα πως οι θεωρητικές εκτιμήσεις συγκλίνουν αρκετά με τις πραγματικές τιμές. Επιπρόσθετα, η εκτίμηση για μεγαλύτερο αριθμό εκτελέσεων βρίσκεται πιο κοντά στην πραγματική όπως άλλωστε αναμενόταν. Ωστόσο ακόμα και μετά από πολύ λίγες εκτελέσεις τα αποτελέσματα είναι αντιπροσωπευτικά. Η διαφορά ανάμεσα στα δύο διαγράμματα είναι ότι στο δεύτερο χρησιμοποιούνται για κάθε εκτέλεση περισσότερα SLAs από το πρώτο οπότε οι μηχανισμοί έχουν καλύτερη εικόνα για τις επιδόσεις τους.

*Η σελίδα αυτή είναι σκόπιμα λευκή*



Το σύστημα από τις πρώτες κιάλας εκτελέσεις μπορούσε να αντιληφθεί τα καλά και τα κακά SLAs, με αποτέλεσμα τα SLA με κακές επιδόσεις να μην επιλέγονται ποτέ και συνεπώς να μην συλλέγονται πληροφορίες για τις επιδόσεις τους.

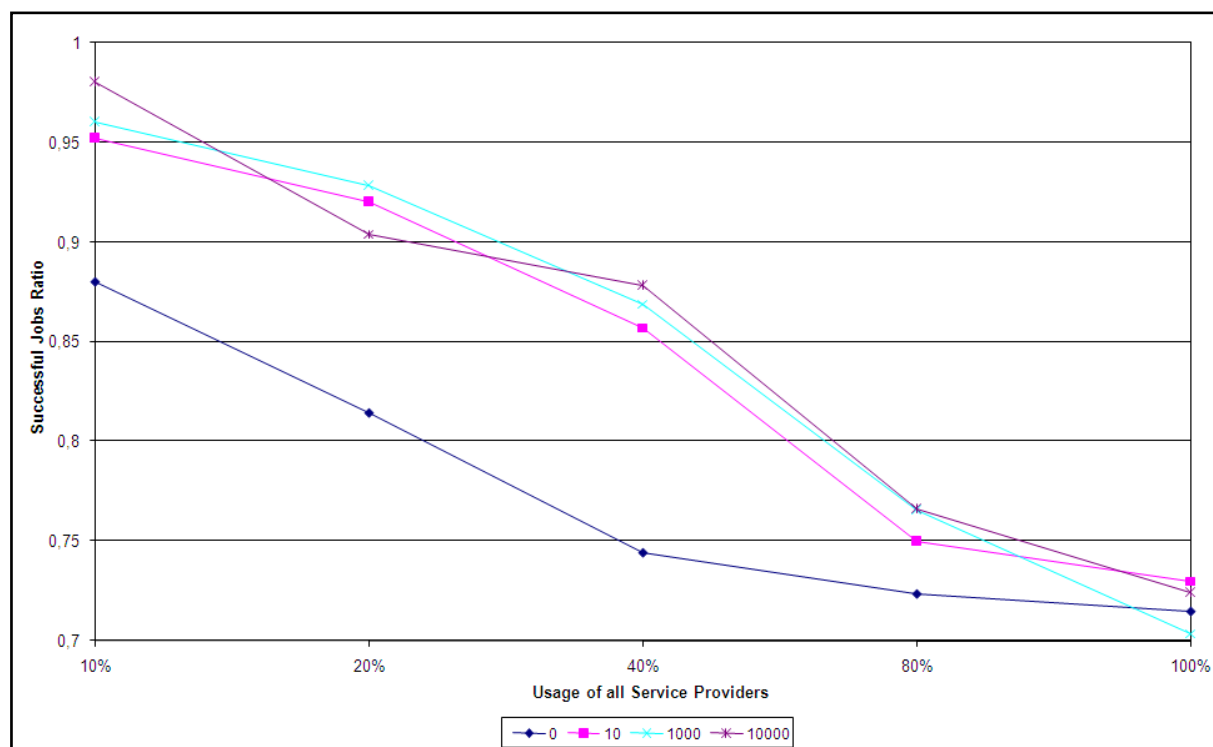
Στο ακόλουθο διάγραμμα (Σχήμα 23) μπορούμε να δούμε την επίδραση που έχει η επιλογή SLAs με τη χρήση μηχανισμών αξιολόγησης στην επιτυχή εκτέλεση εργασιών.



Σχήμα 23: Απόδοση για διαφορετικό αριθμό καταγεγραμμένων δεδομένων

Αρχικά παρατηρούμε ότι χωρίς τη χρήση συστάσεων από τους μηχανισμούς εκτίμησης η επιτυχία εκτέλεσης εργασιών είναι 0.880 για 10% φόρτο εργασίας και 0.744 για φόρτο εργασίας 40%. Καθώς οι μηχανισμοί καταγράφουν περισσότερα δεδομένα η επιτυχία αυξάνεται σε 0.980 για την πρώτη περίπτωση και σε 0.878 για τη δεύτερη. Ωστόσο για μεγαλύτερο φόρτο εργασίας η βελτίωση δεν είναι σημαντική εφόσον τόσο τα SLA με καλές επιδόσεις όσο και τα SLA με κακές πρέπει να επιλεγθούν για να ικανοποιήσουν το φόρτο. Η παρατήρηση αυτή γίνεται περισσότερο ξεκάθαρη στο επόμενο διάγραμμα.

*Η σελίδα αυτή είναι σκόπιμα λευκή*



Σχήμα 24: Απόδοση για διαφορετική χρήση συμφωνιών επιπέδου υπηρεσιών

Αρχικά στο διάγραμμα παρατηρούμε πως γενικά το ποσοστό των επιτυχημένων εργασιών αυξάνεται πολύ σημαντικά χρησιμοποιώντας τους μηχανισμούς αξιολόγησης (χρωματισμένες γραμμές) έναντι της τυχαίας επιλογής (μπλε γραμμή). Εντούτοις, όσο αυξάνεται ο φόρτος εργασίας τόσο η απόδοση των μηχανισμών μειώνεται.

Τα πειραματικά αποτελέσματα έδειξαν σημαντική βελτίωση στην αναλογία επιτυχημένων/αποτυχημένων εργασιών, ωστόσο σε ένα πραγματικό (λειτουργικό) περιβάλλον πλέγματος η απόδοση ήταν ακόμα καλύτερη λόγω του μεγαλύτερου αριθμού δειγμάτων προς αξιολόγηση. Τέλος οι μηχανισμοί αξιολόγησης είναι σε θέση να ακολουθήσουν τις αλλαγές του ποσοστού επιτυχίας και κατά συνέπεια να παρέχουν πάντα την καλύτερη σύσταση για το επίπεδο ποιότητας υπηρεσιών.

### **5.4.3 Πείραμα 2: Δρομολόγηση ροών εργασιών με βάση τις συμφωνίες επιπέδου υπηρεσιών**

Ο σκοπός αυτού του πειράματος ήταν να ερευνηθεί και να αξιολογηθεί η αποτελεσματικότητα της αρχιτεκτονικής πλέγματος και των δομικών συστατικών στο πλαίσιο της εφαρμογής πολυμέσων. Το πείραμα επικεντρώθηκε στα συστατικά που σχετίζονται με τις θεμελιώδεις απαιτήσεις των εφαρμογών πολυμέσων (και γενικά των επιχειρηματικών εφαρμογών) δηλαδή την ποιότητα υπηρεσιών (QoS) και τις συμφωνίες επιπέδου υπηρεσιών (SLAs).

#### **5.4.3.1 Σενάριο**

Το σενάριο που σχεδιάσαμε και υλοποιήσαμε λαμβάνει υπόψη όλες τις καταναλωτικές απαιτήσεις για την διασφάλιση του επιπέδου ποιότητας υπηρεσιών καθώς και τη δυναμική επιλογή ροών εργασιών και υπηρεσιών. Σε αυτό οι καταναλωτές χρησιμοποιούν την υποδομή πλέγματος για την εκτέλεση πραγματικών εργασιών όπως η δημιουργία ολοκληρωμένων κινηματογραφικών σκηνών τρισδιάστατης απεικόνισης. Ο κάθε καταναλωτής υποβάλει τις απαιτήσεις για ποιότητα υπηρεσιών και τα δεδομένα εισόδου μέσω της δικτυακής πύλης η οποία στη συνέχεια αλληλεπιδρά αυτόματα τα απαραίτητα στοιχεία της υποδομής, δημιουργεί τη ροή εργασίας που περιγράφει την εργασία που ο χρήστης επιθυμεί να εκτελέσει βασισμένη πάντα παραμέτρους ποιότητας υπηρεσίας που έχουν ζητηθεί. Στη συνέχεια θα διαπραγματευτεί και θα επιλέξει τα απαραίτητα SLAs και υπηρεσίες και τέλος θα εκτελέσει την εργασία.

#### **5.4.3.2 Αποτελέσματα**

Η πειραματική λειτουργία εφαρμογών πολυμέσων πάνω στη νέα αυτή υποδομή πλέγματος που δημιουργήθηκε με την ενσωμάτωση διαφόρων δομικών στοιχείων έδειξε πως όλα τα δομικά συστατικά και η συνολική υποδομή λειτουργούν αποτελεσματικά και

αρμονικά. Η υποδομή καλύπτει πλήρως τις θεμελιώδεις απαιτήσεις των εφαρμογών πολυμέσων:

- Η δικτυακή πύλη εξασφαλίζει ευκολία στη χρήση και την πρόσβαση των τελικών χρηστών.
- Με την χρήση των μηχανισμών ροών εργασιών και αξιολόγησης των υπηρεσιών αυτοματοποιούνται οι διαδικασίες υποβολής εργασιών
- Με την χρήση των μηχανισμών SLAs και πολιτικών διαχείρισης η υποδομή μπορεί αποδοτικά να διασφαλίζει την παροχή ποιότητας υπηρεσιών.

#### **Οφέλη για τους χρήστες**

Η υποδομή αυτή έχει σημαντικό αντίκτυπο για τους τελικούς χρήστες των εφαρμογών πολυμέσων ή άλλων επιχειρηματικών εφαρμογών με παρόμοιες απαιτήσεις. Η υποδομή προσφέρει στους χρήστες αυτοματοποιημένες διαδικασίες για:

- Δημιουργία και εκτέλεση μια ροής εργασιών
- Αναζήτηση και σύνθεση υπηρεσιών και πόρων
- Επιλογή υπηρεσιών και διαπραγμάτευση συμφωνιών επιπέδου υπηρεσιών με βάση το επίπεδο ποιότητας υπηρεσιών που επιθυμεί ο τελικός χρήστης

Επιπλέον, η υποδομή παρέχει πρόσβαση σε ένα μεγάλο αριθμό πόρων και των υπηρεσιών και κατά συνέπεια η ανάγκη για προμήθεια υλικού εξοπλισμού από την πλευρά του χρήστη εξαλείφεται. Η εξελεγμένη επιλογή υπηρεσιών παρέχει στον τελικό χρήστη ισχυρές εγγυήσεις για την ποιότητα υπηρεσιών ο οποίος μπορεί να χρησιμοποιήσει αποδοτικά τα SLAs. Επίσης οι πολιτικές διαχείρισης καθορίζονται εκ των προτέρων για κάθε SLA ανάλογα με στις ανάγκες της κάθε εφαρμογής. Η αποτυχία ή η καθυστέρηση στη λήψη του αποτελέσματος αντιμετωπίζεται με τα εξελεγμένα SLAs με τον ορισμό συγκεκριμένων αποζημιώσεων. Τέλος τα τμήματα λειτουργικής και επιχειρησιακής

διαχείρισης παρέχουν τον προηγμένο χειρισμό των εργασιών και κατά συνέπεια αυξάνουν το κέρδος των παρόχων προσφέροντας ποιοτικότερες υπηρεσίες στους πελάτες τους.

# 6

## *Λειτουργία του πλέγματος σε πραγματικό χρόνο*

Οι υπηρεσιοστρεφείς αρχιτεκτονικές (Service Oriented Infrastructures – SOA) θεωρούνται όλο και περισσότερο ως ιδανικές επιλογές για την υλοποίηση υποδομών που μπορούν να υποστηρίξουν τις επιχειρησιακές εφαρμογές. Επίσης, μια βασική αρχή λειτουργίας τέτοιων επιχειρησιακών περιβαλλόντων είναι ότι οι σχέσεις μεταξύ παρόχων και χρηστών υπηρεσιών καθορίζονται από τα SLA (Service Level Agreement) [21],[12]. Αυτό σημαίνει ότι οι ρόλοι των συμβαλλόμενων μερών που συμμετέχουν σε μια επιχειρησιακή συναλλαγή σε ένα περιβάλλον SOA (πελάτης φορέων παροχής υπηρεσιών και υπηρεσιών), η ίδια η σχέση και οι λεπτομέρειες της σχέσης καθορίζονται από SLA. Αυτές οι λεπτομέρειες διευρύνονται σε ολόκληρη την έκταση του κύκλου ζωής της σχέσης όπως οι χρόνοι εγκατάστασης και παροπλισμού και η φύση της σχέσης τους.

Αρχικά και βασισμένοι σε πρότυπα όπως το WS-Agreement [101], τα SLA περιελάμβαναν μόνο τις πληροφορίες για την ποιότητα μιας υπηρεσίας κατά τη διάρκεια της διατήρησης, της έναρξης και της χρήσης της. Σε μερικές περιπτώσεις το κόστος χρήσης του συγκεκριμένου επιπέδου υπηρεσίας (QoS) για την υπηρεσία συμπεριλαμβάνεται επίσης στα SLA. Εντούτοις, διάφορα σύγχρονα επιχειρησιακά περιβάλλοντα απαιτούν περιπλοκότερα σχήματα SLA [64] για να υποστηρίξουν: α) την παροχή πρόσθετων πληροφοριών σχετικά με τη χρήση υπηρεσιών (π.χ. ανά πόρο ή ανά

δραστηριότητα & στιγμιαία ή συσσωρευτικά) β) χειρισμός των παραβιάσεων SLA που συνδέονται με την επιβολή των αποζημιώσεων γ) μηχανισμοί για τον έλεγχο SLA και, δ) κανόνες για την επιβολή SLA και τη συντήρηση του συμφωνηθέντος επιπέδου QoS. Αυτή η περίπλοκη προσέγγιση στα SLA απαιτεί συνδέσεις με τις συγκεκριμένες πολιτικές και τις δηλώσεις στο SLA για το ποιά πολιτική πρέπει να χρησιμοποιηθεί όταν επιτυγχάνεται το κατώτατο όριο μιας παραμέτρου. Αυτό το κατώτατο όριο είναι ιδιαίτερης σπουδαιότητας δεδομένου ότι δείχνει ότι μια παραβίαση είναι έτοιμη να εμφανιστεί. Οι προαναφερθείσες πολιτικές είναι συγκεκριμένες ανά εφαρμογή (δεδομένου ότι αναφέρονται στους ιδιαίτερους όρους του SLA) καθορίζοντας τους ρητούς κανόνες για τις κρίσιμες αποφάσεις που πρέπει να ληφθούν ώστε να εγγυηθεί το απαραίτητο επίπεδο QoS για κάθε όρο του SLA, λαμβάνοντας υπόψη διάφορες πληροφορίες όπως οι γενικές απαιτήσεις QoS, το κόστος, και τις πληροφορίες ελέγχου σε πραγματικό χρόνο.

Τα SLA και συνεπώς οι πολιτικές, πρέπει να καθορίσουν ρητά τους κανόνες και τις ενέργειες που απαιτούνται για τη διαδικασία παροχής QoS [4], ειδικά στα επιχειρησιακά περιβάλλοντα πραγματικού χρόνου. Οι αποφάσεις κατά τη διάρκεια αυτής της διαδικασίας απεικονίζονται στη γενική λειτουργία και την απόδοση του συστήματος, δεδομένου ότι οι υπηρεσίες και οι παράμετροι QoS για κάθε υπηρεσία επανυπολογίζονται σε ένα σύντομο χρονικό πλαίσιο λόγω των περιορισμών της εφαρμογής που εκτελείται σε πραγματικό χρόνο. Ένα παράδειγμα αυτού του τύπου εφαρμογών είναι η τρισδιάστατη επεξεργασία [20], η οποία χρησιμοποιείται επίσης προκειμένου να αξιολογηθεί η προτεινόμενη μέθοδος παροχής QoS.

Παρακάτω παρουσιάζουμε μια μέθοδο παροχής QoS [8] που είναι βασισμένη όχι μόνο στα στοιχεία ελέγχου πραγματικού χρόνου, αλλά που επίσης εκμεταλλεύεται τις πληροφορίες από τις προηγούμενες εκτελέσεις. Επιπλέον, είναι εξατομικεύσιμο για κάθε εφαρμογή, δεδομένου ότι χρησιμοποιεί συγκεκριμένες πολιτικές εφαρμογής. Αυτή η



προσέγγιση λαμβάνει υπόψη όλες τις παραμέτρους που περιλαμβάνονται στα σχήματα SLA, όπως οι όροι αποζημιώσεων, και με αυτό τον τρόπο προτείνει μια βέλτιστη λύση από την άποψη της απόδοσης QoS που είναι επίσης οικονομικά δίκαιη για το φορέα παροχής υπηρεσιών.

Η παροχή QoS και η διαχείριση του γενικά είναι ένας πολύ σημαντικός και προκλητικός τομέας της έρευνας στις υπηρεσιοστρεφείς υποδομές. Στο [39] παρουσιάζεται μια προσέγγιση για τη διαχείριση QoS σε SOA που εστιάζει στην κατηγοριοποίηση των ποιοτικών χαρακτηριστικών και χρησιμοποιεί μια συγκεκριμένη γλώσσα, βασισμένη σε XML, για τις εφαρμογές και τους φορείς παροχής υπηρεσιών για να εκφράσει τις απαιτήσεις του QoS και των συμβάσεων. Για τις ελαφριές σε πραγματικό χρόνο εφαρμογές, στις οποίες η χρησιμοποίηση των πόρων κάθε εργασίας είναι σημαντικό να μείνει υπό έλεγχο, οι συντάκτες στο [75] προτείνουν τον συνδυασμό ενός χρονοπρογραμματιστή διατήρησης των πόρων και έναν μηχανισμό ανατροφοδότησης για την δυναμική προσαρμογή του κλάσματος της ΚΜΕ (εύρος ζώνης) που διατίθεται στις εργασίες. Παρόμοιος με αυτόν, ένας μηχανισμός για στις εφαρμογές πραγματικού χρόνου που χρησιμοποιεί προσαρμοστικούς αλγορίθμους διατήρησης και ελέγχου που εμπνέονται από τις ιδέες του στοχαστικού ελέγχου περιγράφεται μέσα στα [61] και [60]. Επιπλέον, η δημοσίευση [36] επεξηγεί ένα πρότυπο για την παροχή QoS των εφαρμογών πολυμέσων, το οποίο είναι βασισμένο σε κάποια μορφή διατήρησης των πόρων από κοινού με την υποστήριξη επιπέδων συστημάτων για τις προσαρμοστικές εφαρμογές. Τέλος, το [47] περιγράφει ένα πλαίσιο για τον έλεγχο και την ανάλυση των μετρικών QoS προκειμένου να εγγυηθεί το επίπεδο QoS. Σε αυτήν την προσέγγιση οι πόροι, οι υπηρεσίες και οι εξαρτήσεις τους διαμορφώνονται και χρησιμοποιούνται διάφορες τεχνικές για να αναλύσουν τις μετρικές QoS των εξαρτώμενων υπηρεσιών πλέγματος. Η διαφορά αυτών των προσεγγίσεων σε σύγκριση με την προτεινόμενη βρίσκεται στο γεγονός ότι η

προσέγγισή μας χειρίζεται QoS όχι μόνο στο επίπεδο των πόρων (π.χ. [39]) αλλά είναι καθοδηγούμενη από το SLA και πλήρως αξιοποιήσιμη χρησιμοποιώντας συγκεκριμένες πολιτικές εφαρμογής. Επιπλέον, οι ικανότητες πραγματικού χρόνου της παρουσιαζόμενης προσέγγισης, ενισχύονται με μια διαδικασία αξιολόγησης του γεγονότος και έναν μηχανισμό σύστασης που εκμεταλλεύεται τα ιστορικά αρχεία κάθε εργασίας.

### ***6.1 Παροχή ποιότητας υπηρεσιών σε υποδομές πραγματικού χρόνου***

Προκειμένου να εγγυηθεί το συμφωνηθέν επίπεδο QoS στα πραγματικού χρόνου περιβάλλοντα, η κατάσταση των πόρων και των εκτελέσεων εργασιών πρέπει να ελέγχονται συνεχώς όπως επίσης και οι τιμές των παραμέτρων QoS. Με βάση αυτές τις πληροφορίες, μπορεί να απαιτηθεί ο επανασηματισμός εργασίας προκειμένου να βεβαιωθεί το απαραίτητο επίπεδο QoS. Συνήθως ο αριθμός των ελεγχόμενων παραμέτρων είναι αρκετά υψηλός και εξαρτάται από το πόσο περίπλοκη είναι μια εργασία. Για τις εφαρμογές με υψηλές απαιτήσεις QoS, όπως η τρισδιάστατη επεξεργασία [20], πρέπει να ελεγχθούν περισσότερες παράμετροι απ' ό,τι για τις μάλλον τετριμμένες εφαρμογές, από την άποψη των απαιτήσεων QoS, όπως η τηλεδιάσκεψη. Επιπλέον, το σύνολο των παραμέτρων ποικίλλει από μια εφαρμογή σε μια άλλη και μια παράμετρος που έχει επιπτώσεις στο QoS μιας εφαρμογής μπορεί να είναι αδιάφορη για κάποια άλλη.

Στην προτεινόμενη μέθοδο παροχής QoS, μια διαδικασία ιδιαίτερης σπουδαιότητας είναι η αξιολόγηση των ελεγχόμενων παραμέτρων. Η διαδικασία αξιολόγησης είναι απλή για τις παραμέτρους εφαρμογής που συχνά ταιριάζουν με τους όρους SLA αλλά είναι αρκετά περίπλοκη για τις παραμέτρους ελέγχου χαμηλών επιπέδων όπου τις περισσότερες φορές πρέπει να αντιστοιχισθούν στις παραμέτρους υψηλού επιπέδου. Σε αυτό το έγγραφο ακολουθήσαμε τις ίδιες μεθόδους αντιστοίχισης QoS για τη διαδικασία παροχής QoS που χρησιμοποιούνται επίσης κατά τη διάρκεια των φάσεων ανακαλύψεων και

διαπραγμάτευσης SLA και παρουσιάζονται εδώ [58]. Με αυτό τον τρόπο, η διαδικασία αξιολόγησης, που είναι αξιοποιήσιμη για κάθε εφαρμογή που χρησιμοποιεί τις πολιτικές αντιστοίχισης, συνδέει τα γεγονότα και τις παραμέτρους ελέγχου με τις ποιοτικές παραμέτρους υψηλού επιπέδου και είναι σε θέση να προσδιορίσει την σημαντικότερο από αυτές.

Εντούτοις, το σύστημα και ο επανασηματισμός εκτέλεσης εργασιών για την παροχή QoS είναι πολύ περίπλοκοι για τις πραγματικού χρόνου εφαρμογές, ακόμη και με την αξιολόγηση των παραμέτρων ελέγχου και τη χαρτογράφηση των συγκεκριμένων για την εφαρμογή, δεδομένου ότι τα μέρη των στοιχείων πρέπει να αναλυθούν σε ένα σύντομο χρονικό πλαίσιο προκειμένου να ληφθεί μια σωστή απόφαση και να προβλεφθεί η κατάσταση συστημάτων μετά από τον επανασηματισμό. Αυτό μπορεί τελικά να οδηγήσει στις παραβιάσεις SLA και τα μειωμένα επίπεδα της προσφερόμενης ποιότητας.

Δεδομένου ότι οι αποφάσεις πρέπει να ληφθούν σε ένα σύντομο χρονικό πλαίσιο, μια καινοτομία του προτεινόμενου πλαισίου είναι να υποστηριχθεί η διαδικασία παροχής QoS με έναν μηχανισμό σύστασης. Η έννοια της σύστασης είναι βασισμένη στη χρήση των ιστορικών πληροφοριών για την αξιολόγηση των ικανοτήτων ποιοτικής παροχής ενός φορέα παροχής υπηρεσιών. Σε όλο το χρόνο και τη χρήση, διάφορα γεγονότα καθορίζουν ένα συμπεριφοριστικό σχεδιάγραμμα ενός φορέα παροχής υπηρεσιών. Ελέγχοντας επιλεκτικά μερικά γνωρίσματα που χαρακτηρίζουν αυτήν την συμπεριφορά, ο φορέας παροχής υπηρεσιών μπορεί να κάνει εκτιμήσεις στο τι να περιμένει (και να αποφασίσει πώς να αντιδράσει) στο μέλλον υπό τους συγκεκριμένους όρους. Με υψηλή πιθανότητα, η αξιοπιστία αυτού του μηχανισμού βελτιώνεται καθώς συλλέγονται περισσότερα στοιχεία σχετικά με τα ελεγχόμενα χαρακτηριστικά γνωρίσματα. Στη λειτουργία ενός φορέα παροχής υπηρεσιών είναι μάλλον ασφαλές να υποτεθεί ότι οι ιστορικές πληροφορίες θα

αναφερθούν συχνά δεδομένου ότι αυτές οι πληροφορίες μπορούν να παραχθούν από κάθε ενέργεια.

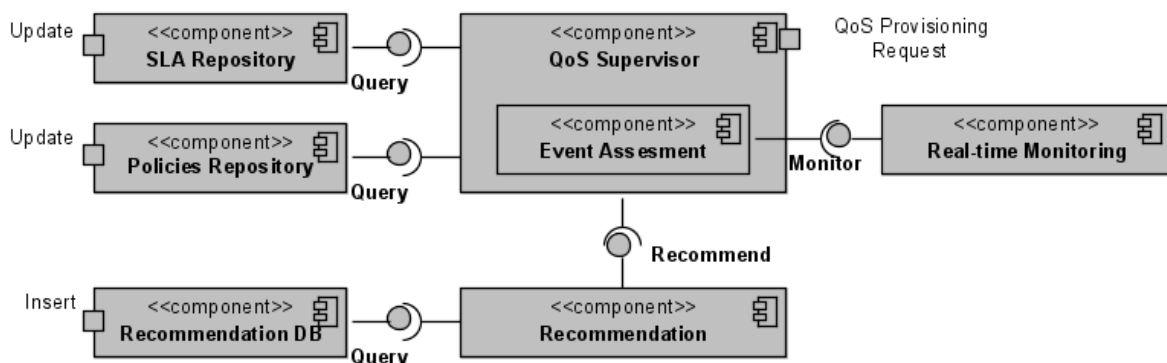
Με βάση τα ανωτέρω, αυτό που είναι σημαντικό είναι το σύνολο των χαρακτηριστικών γνωρισμάτων που διέπουν τη συμπεριφορά ενός φορέα παροχής υπηρεσιών. Δεν υπάρχει κανένα γενικό πρότυπο για να καθορίσει τέτοια χαρακτηριστικά γνωρίσματα όπως επηρεάζονται από το πλαίσιο της εφαρμογής καθώς επίσης και από το συγκεκριμένο πελάτη. Οποιοσδήποτε σχεδιαστής τρισδιάστατων γραφικών μπορεί να έχει διαφορετικά κριτήρια για να αξιολογήσει την ποιότητα μιας εργασίας και ακόμα κι αν αυτά τα κριτήρια επικαλύπτονται περιστασιακά, είναι σίγουρα διαφορετικά από τις άλλες εφαρμογές. Επομένως ο μηχανισμός σύστασης πρέπει να είναι αρκετά προσαρμοστικός να υποστηρίξει οποιοδήποτε τύπο κριτηρίων εφαρμογής και - στην ουσία - ποιότητας. Ο πελάτης πρέπει να καθορίσει τα κριτήρια και στη συνέχεια ο υπεύθυνος για την ανάπτυξη εφαρμογής πρέπει να προσδιορίσει τις πηγές και τους μηχανισμούς ελέγχου που θα παράσχουν τα στοιχεία σχετικά με τα κριτήρια, κάθε φορά που καλείται μια υπηρεσία. Κατά συνέπεια, ο φορέας παροχής υπηρεσιών πρέπει πρώτα να διαμορφώσει το μηχανισμό σύστασης για να προσκομίσει τα στοιχεία από τις αντίστοιχες πηγές μετά από τα συγκεκριμένα γεγονότα και έπειτα να διαμορφώσει το μηχανισμό για να καταλάβει πώς να αξιολογήσει αυτά τα στοιχεία μαζικά μετά από άλλο τύπο γεγονότων. Ο πρώτος τύπος γεγονότων παράγεται όποτε καλείται μια υπηρεσία (κατά τη διάρκεια της εκτέλεσης) ενώ ο δεύτερος όποτε μια υπηρεσία πρόκειται να επιλεγεί.

Η προτεινόμενη μέθοδος παροχής QoS αξιοποιεί τις προαναφερθείσες διαδικασίες, της αξιολόγησης του γεγονότος και της σύστασης βασισμένων στα ιστορικά στοιχεία. Μέσα σε αυτήν την μέθοδο, ο μηχανισμός σύστασης χρησιμοποιεί την προηγούμενη εμπειρία για να εκτιμήσει τις ικανότητες ποιοτικής παροχής υπηρεσιών και να συστήσει την καλύτερη πολιτική που πρέπει να εφαρμοστεί για να εγγυηθεί το επίπεδο QoS. Για κάθε

πολιτική η σύσταση απεικονίζει πόσο καλά συμπεριφέρθηκε η συγκεκριμένη πολιτική στο παρελθόν υπό τους ίδιους ή παρόμοιους όρους. Με αυτόν τον τρόπο η παρουσιασμένη μέθοδος παροχής QoS, προκειμένου να εγγυηθεί το ποιοτικό επίπεδο, προτείνει τον επανασχηματισμό του συστήματος που καθορίζεται στην κατάλληλη, για τους τρέχοντες όρους, πολιτική γνωρίζοντας πώς αυτός ο επανασχηματισμός έχει εκτελεσθεί στο παρελθόν και εάν εμφανίστηκαν παραβιάσεις SLA.

## 6.2 Αρχιτεκτονική προσέγγιση

Η έννοια που περιγράφεται πιο πάνω απαιτεί το σχεδιασμό και την ενσωμάτωση των συγκεκριμένων αρχιτεκτονικών στοιχείων, στο πεδίο των φορέων παροχής υπηρεσιών, προκειμένου να παρασχεθεί αυτή η λειτουργία για τα πραγματικού χρόνου πλέγματα. Σε αυτό το τμήμα, περιγράφουμε μια υλοποίηση αναφοράς που ακολουθεί όλες τις αρχές SOA [13] και επομένως ισχύει σε οποιοδήποτε SOA. Όλα τα αρχιτεκτονικά στοιχεία σε αυτό το πλαίσιο εκθέτουν τη λειτουργία τους και επικοινωνούν το ένα το άλλο μέσω των συγκεκριμένων διεπαφών. Αφ' ετέρου, η αλληλεπίδραση με τα άλλα δομικά στοιχεία της υποδομής εκτελείται χρησιμοποιώντας εξωτερικές υποδοχές. Ο κύκλος ζωής του SLA όπως περιγράφεται στο [77] περιλαμβάνει τις συγκεκριμένες διαδικασίες για τη διαχείριση SLA που ακολουθήσαμε ακριβώς για το γενικό σχέδιο πλαισίου.



Σχήμα 25: Διάγραμμα Αρχιτεκτονικής

*Η σελίδα αυτή είναι σκόπιμα λευκή*

Το προηγούμενο διάγραμμα (component diagram) απεικονίζει τη γενική αρχιτεκτονική του πλαισίου παροχής QoS και όλων των δομικών στοιχείων, των διεπαφών και των θυρών που αποτελούν αυτό το πλαίσιο. Το κύριο συστατικό του προτεινόμενου πλαισίου (όπως απεικονίζεται στο διάγραμμα ανωτέρω) είναι ο επόπτης QoS που επικοινωνεί με όλα τα άλλα συστατικά και περιλαμβάνει μια υποδοχή στην υποδομή για να αρχίσει το αίτημα παροχής QoS. Μετά από, περιγράφουμε αυτά τα συστατικά μαζί με τις λεπτομέρειες της εκτέλεσής τους.

### **6.2.1 Αποθήκη συμφωνιών επιπέδου υπηρεσιών**

Το στοιχείο αυτό χρησιμοποιείται για να αποθηκεύσει τα SLA για το συγκεκριμένο φορέα παροχής υπηρεσιών. Κάθε SLA είναι μοναδικό, και η αποθήκη κρατά όχι μόνο τη "μοναδική ταυτότητα" (unique identifier) του SLA και τα περιληφθέντα συμβαλλόμενα μέρη αλλά και όλες τις πληροφορίες του SLA, όπως οι όροι SLA, η τιμή και οι αποζημιώσεις. Ο επόπτης QoS επικοινωνεί με την αποθήκη SLA χρησιμοποιώντας τη διεπαφή ερώτησης προκειμένου να ληφθούν οι λεπτομέρειες SLA που περιλαμβάνουν τις ποιοτικές πληροφορίες για κάθε εργασία που τρέχει κάτω από το συγκεκριμένο SLA. Επιπλέον, η αποθήκη SLA περιλαμβάνει έναν υποδοχέα για την ενημέρωση των SLA που αποθηκεύονται σε αυτό. Πρέπει να αναφερθεί ότι λόγω του ότι τα SLA είναι μοναδικά είναι βασισμένα στα συγκεκριμένα πρότυπα, οι λεπτομέρειες SLA δεν μπορούν να αλλάξουν κατά τη διάρκεια του χρόνου εκτέλεσης.

### **6.2.2 Αποθήκη πολιτικών διαχείρισης**

Στην προτεινόμενη εφαρμογή αναφοράς κάθε SLA συνδέεται με τις πολιτικές που περιλαμβάνουν τους διοικητικούς κανόνες είτε για το γενικό SLA είτε για τους ιδιαίτερους όρους SLA. Αυτοί οι κανόνες καθορίζονται από το φορέα παροχής υπηρεσιών και εφαρμόζονται ώστε να εγγυηθεί το απαραίτητο επίπεδο QoS ή να εξεταστούν οι συγκεκριμένες απαιτήσεις χρόνου εκτέλεσης κάθε εφαρμογής. Η αποθήκη πολιτικών

αποθηκεύει τις πολιτικές και τις συνδέει με SLA ή με τους ιδιαίτερους όρους SLA. Ο φορέας παροχής υπηρεσιών είναι σε θέση να ενημερώσει τις πολιτικές κατά τη διάρκεια του χρόνου εκτέλεσης, να επεκτείνει νέες ή να αλλάξει τις συσχετίσεις μέσω του υποδοχέα αναπροσαρμογών.

### **6.2.3 Έλεγχος πραγματικού χρόνου**

Το κομμάτι του τμήματος *ελέγχου πραγματικού χρόνου* συλλέγει σε πραγματικό χρόνο πληροφορίες από κάθε πόρο του φορέα παροχής υπηρεσιών. Αυτό το στοιχείο συσχετίζεται όχι μόνο με μια συγκεκριμένη εργασία αλλά με όλες τις υπηρεσίες και τους πόρους του φορέα παροχής υπηρεσιών. Γενικά, οι μετρικές για τις παραμέτρους διαβιβάζονται στο τμήμα *αξιολόγησης του γεγονότος* μέσω της διεπαφής *οργάνων ελέγχου* και αξιολογούνται έπειτα για τη σχέση τους με την εργασία και για τις επιπτώσεις τους στο επίπεδο QoS. Φυσικά οι πληροφορίες ελέγχου που συλλέγονται από το *τμήμα ελέγχου πραγματικού χρόνου* εξαρτώνται από την εφαρμογή και πώς αυτή προσαρμόζεται ως υπηρεσία στη συγκεκριμένη υπηρεσιοστρεφή υποδομή (SOI). Τα γεγονότα ελέγχου ακολουθούν μια ιεραρχία από τα γενικά στα γεγονότα της συγκεκριμένης εφαρμογής επομένως οι πολιτικές που συνδέονται με τα γεγονότα αυτά ακολουθούν μια παρόμοια ιεραρχία. Για καθένα αυτών των γεγονότων ο *επόπτης QoS* προκαλεί την πιο κατάλληλη πολιτική (οι συγκεκριμένες πολιτικές εφαρμογής επιλέγονται πρώτα βασισμένες στην προαναφερθείσα ιεραρχία) που περιλαμβάνει τους "καλύτερους" κανόνες, από άποψη απόδοσης, για αυτό το σύνολο γεγονότων.

Το *τμήμα ελέγχου πραγματικού χρόνου* μπορεί να χρησιμοποιήσει διαφορετικές προσεγγίσεις, όπως η προώθηση (push) ή απόσυρση (pull) ανακοίνωσης, για τη συλλογή των πληροφοριών σε πραγματικό χρόνο αλλά οι λεπτομέρειες σχετικά με αυτό είναι έξω από το πεδίο ερευνάς μας. Στην εφαρμογή μας χρησιμοποιήθηκε η μέθοδος ανακοίνωσης



προώθησης (push notification), συναθροίζοντας τα γεγονότα πραγματικού χρόνου από τα wrappers-scripts εκτέλεσης της εφαρμογής και αναμεταδίδοντας τα στον *επόπτη QoS*.

#### **6.2.4 Τμήματα σύστασης**

Η διαδικασία παροχής QoS ενισχύεται χρησιμοποιώντας έναν μηχανισμό σύστασης. Αυτός ο μηχανισμός αποτελείται από δύο τμήματα, τη *σύσταση* και τη *βάση δεδομένων σύστασης* (ΒΔ). Η ΒΔ *σύστασης* αποθηκεύει τις ιστορικές πληροφορίες από όλες τις προηγούμενες εκτελέσεις εργασίας. Συγκεκριμένα, αυτό το συστατικό καταγράφει όλα τα στοιχεία για κάθε εργασία όπως ο τύπος εργασίας, πόροι που χρησιμοποιήθηκαν, οι παραβιάσεις SLA που εμφανίστηκαν και όλα αυτά ταξινομήθηκαν ώστε να εξαχθούν εύκολα. Επίσης όλες οι προηγούμενες αποφάσεις του *επόπτη QoS* και των πολιτικών που χρησιμοποιήθηκαν καταγράφονται και όλες συνδέονται με τον τύπο εργασίας και τους όρους SLA QoS. Ο ρόλος του *τμήματος σύστασης* είναι να αξιολογηθούν τα ιστορικά στοιχεία της ΒΔ *σύστασης* βασισμένης στην τρέχουσα θέση επιπέδων QoS και τις πληροφορίες ελέγχου και να παρασχεθεί ένα μέτρο για κάθε λύση για το πώς το περιβάλλον αποκρίθηκε στο παρελθόν υπό περίπου τους ίδιους όρους. Η επικοινωνία των δύο συστατικών εκτελείται μέσω της διεπαφής *ερώτησης της ΒΔ σύστασης* και ολόκληρη η διαδικασία σύστασης προκαλείται μέσω της διεπαφής *σύστασης*.

#### **6.2.5 Επόπτης ποιότητας υπηρεσιών**

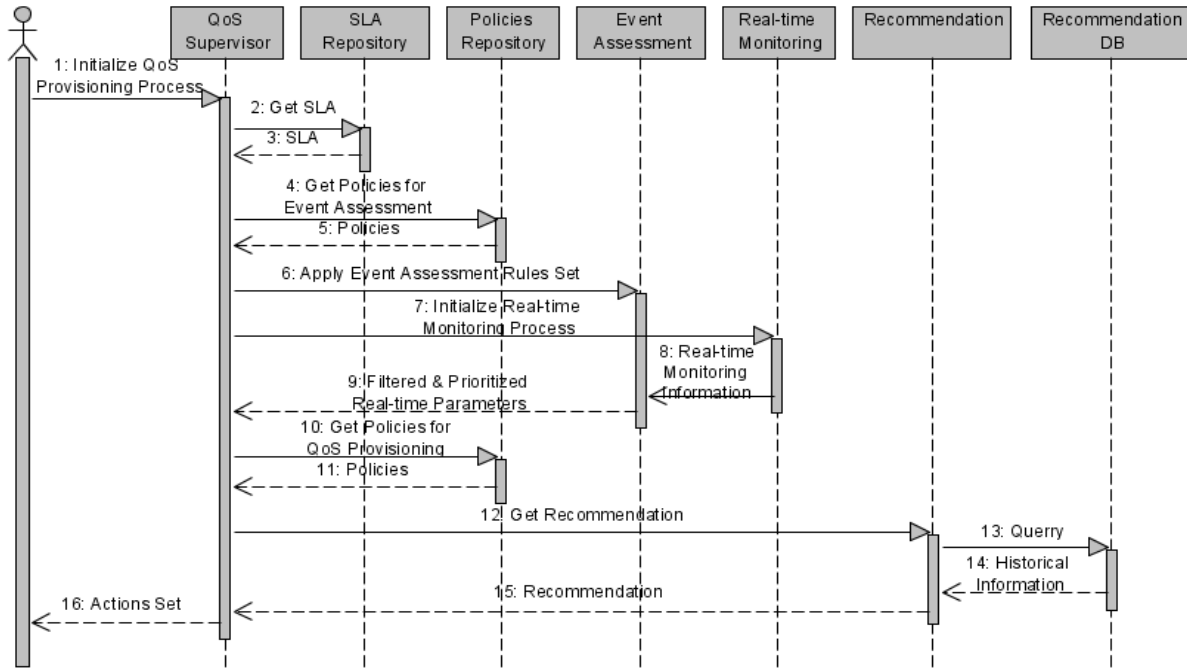
Όπως αναφέρθηκε ήδη ο *επόπτης QoS* είναι το κύριο συστατικό του προτεινόμενου πλαισίου και αρχικοποιεί τις διαδικασίες παροχής QoS για κάθε εργασία. Το αίτημα παροχής QoS προκαλείται, χρησιμοποιώντας τον αντίστοιχο υποδοχέα, από το υλικολογισμικό ή οποιαδήποτε άλλη υπηρεσία στην περιοχή φορέων παροχής υπηρεσιών ανάλογα με το πώς αυτός ο μηχανισμός είναι ενσωματωμένος στη γενική υποδομή. Αυτό το συστατικό επικοινωνεί με όλα τα προαναφερθέντα συστατικά μέσω των διεπαφών τους και αποκρίνεται με μια συγκεκριμένη απόφαση σχετικά με τις παραμέτρους και τους

πόρους που πρέπει να ρυθμιστούν ώστε να διατηρηθούν το ποιοτικό επίπεδο για κάθε όρο SLA ή τον υψηλότερο δυνατό για το συνολικό SLA, δεδομένου ότι κάθε SLA είναι μοναδικό και δεν μπορεί να αλλάξει. Αυτή η λειτουργία εκτελείται μέσω της διεπαφής *ερώτησης* του τμήματος *αποθηκών SLA* και αυτές οι πληροφορίες αποθηκεύονται προσωρινά μέσα στον *επόπτη QoS* κατά τη διάρκεια του χρόνου εκτέλεσης μιας εργασίας ώστε να είναι αμέσως διαθέσιμες. Αφ' ετέρου οι πολιτικές ανακτώνται από την *αποθήκη πολιτικών* κάθε φορά που απαιτούνται, δεδομένου ότι μπορούν να ενημερωθούν δυναμικά.

Η *Αξιολόγηση του γεγονότος* είναι ένα υπο-τμήμα του *επόπτη QoS* που φιλτράρει και βάζει σε προτεραιότητα τις μετρικές πραγματικού χρόνου ώστε να αξιοποιηθούν περισσότερο αποτελεσματικά από τον *επόπτη QoS*. Η διαδικασία της αξιολόγησης είναι πολύ σημαντική καθότι επηρεάζει σημαντικά την απόδοση πραγματικού χρόνου του πλαισίου εργασίας και φυσικά το αποτέλεσμα της παροχής QoS στην περίπτωση που οι παράμετροι δεν έχουν μπει σε σωστή προτεραιότητα ή δεν είναι συνδεδεμένοι με τους σωστούς όρους SLA. Αυτό το τμήμα έχει πλήρη διασύνδεση και συγκεκριμένες πολιτικές εφαρμόζονται για κάθε εφαρμογή.

### **6.3 Αλληλεπιδράσεις βασικών δομικών στοιχείων**

Ο κύριος στόχος της διαδικασίας παροχής QoS είναι να καθοριστούν οι κανόνες που προκύπτουν από τις πολιτικές και που πρέπει αυτές να εφαρμοστούν προκειμένου να εγγυηθεί το ζητούμενο ποιοτικό επίπεδο. Η ροή των πληροφοριών μεταξύ των συστατικών που πραγματοποιούν αυτήν την διαδικασία απεικονίζεται στον ακόλουθο διάγραμμα (Σχήμα 26).



Σχήμα 26: Διάγραμμα Αλληλεπίδρασης

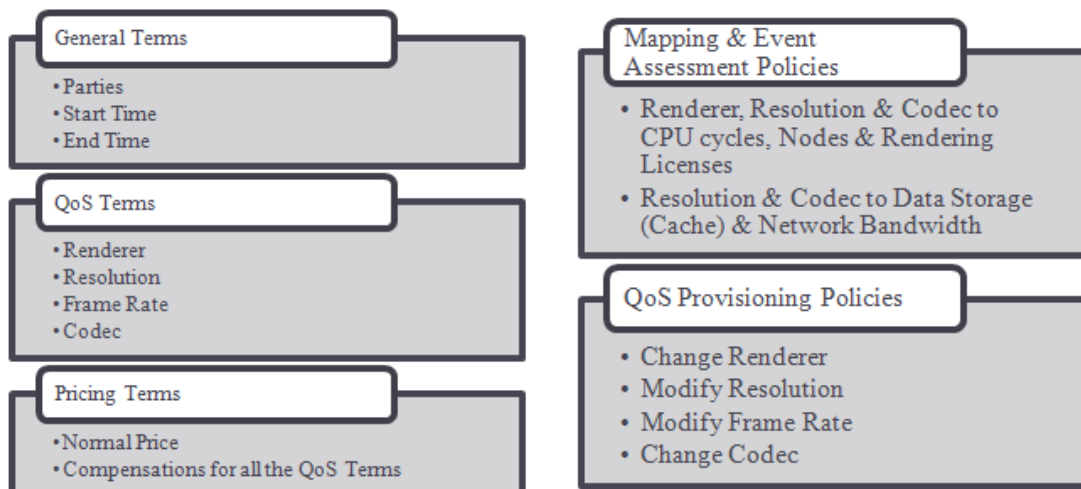
Η διαδικασία παροχής QoS αρχίζει, όταν ένας δράστης (μια υπηρεσία του φορέα παροχής υπηρεσιών ή του υλικολογισμικού) ζητά την παροχή μιας συγκεκριμένης εργασίας από το QoS επόπτη (1). Ο επόπτης QoS έρχεται σε επαφή με την αποθήκη SLA για να λάβει το SLA που συνδέεται με τη συγκεκριμένη εργασία και συντηρεί ένα αντίγραφο από αυτήν (2, 3). Το επόμενο βήμα είναι να αποκτηθούν οι πολιτικές που συσχετίζονται με την αξιολόγηση του γεγονότος από την αποθήκη πολιτικών (4, 5) στη συνέχεια να αναπτυχθεί το σύνολο κανόνων στην Αξιολόγηση του Γεγονότος (6) και η αρχικοποίηση του ελεγκτή πραγματικού χρόνου (7). Το τμήμα ελέγχου πραγματικού χρόνου υποβάλλει συνεχώς τις real-time πληροφορίες (Η ασύγχρονη δράση 8) και η Αξιολόγηση του Γεγονότος τις αναλύει και τις συνδέει με τα συγκεκριμένα στοιχεία του SLA. Ένα πλήρες σύνολο πλέον σημαντικών παραμέτρων αναμεταδίδεται έπειτα στο QoS επόπτη (9). Βασισμένος στις καθορισμένες παραμέτρους, ο επόπτης QoS λαμβάνει τις αντίστοιχες πολιτικές (10, 11) και ζητά έκτοτε μια σύσταση (12). Η είσοδος για την ερώτηση (query) σύστασης (13, 14) είναι όχι μόνο οι διαθέσιμες πολιτικές, αλλά και οι όροι SLA και οι παράμετροι ελέγχου

*Η σελίδα αυτή είναι σκόπιμα λευκή*

που αντιπροσωπεύουν πραγματικά τους τρέχοντες συνθήκες QoS. Μια σύσταση επιστρέφεται στο QoS επόπτη (15), ο οποίος παράγει τελικά ένα συγκεκριμένο σύνολο ενεργειών που θα διατηρούσε το επίπεδο QoS ή θα εγγυάται το πιο υψηλό γενικό επίπεδο QoS για την εργασία υπό τις ορισμένες συνθήκες.

#### 6.4 Το σενάριο της τρισδιάστατης απεικόνισης

Η προτεινόμενη ιδέα παροχής QoS και η προαναφερθείσα υλοποίηση αναφοράς αναμένονται να είναι εφαρμόσιμες για όλες τις εφαρμογές πραγματικού χρόνου, σε οποιοδήποτε SOA. Αυτή η εφαρμογή επικυρώθηκε μέσω ενός πραγματικού επιχειρησιακού σεναρίου, για την εφαρμογή τρισδιάστατης απεικόνισης [20]. Αυτό το σενάριο περιλαμβάνει έναν σχεδιαστή τρισδιάστατων κινούμενων εικόνων που έχει πρόσβαση σε ένα περιβάλλον πλέγματος για να υποβάλει μια σκηνή wireframe σε έναν φορέα παροχής υπηρεσιών προς επεξεργασία. Η επεξεργασμένη σκηνή επιστρέφεται στο σχεδιαστή ως video stream. Η επιχειρησιακή σχέση μεταξύ αυτών των συμβαλλόμενων μερών καθιερώνεται κάτω από ένα SLA και επίσης ένα σύνολο πολιτικών που σχεδιάζονται από τον υπεύθυνο για την ανάπτυξη εφαρμογής ή το φορέα παροχής υπηρεσιών που συνδέεται με αυτό το SLA.



Σχήμα 27: Συμφωνίες επιπέδου υπηρεσιών και πολιτικές διαχείρισης για την εφαρμογή 3D Rendering

*Η σελίδα αυτή είναι σκόπιμα λευκή*

Το προηγούμενο διάγραμμα απεικονίζει μια επισκόπηση των όρων SLA και των πολιτικών για την τρισδιάστατη απεικόνιση. Από την πλευρά φορέων παροχής υπηρεσιών η *ΒΔ σύστασης* περιλαμβάνει ήδη τα ιστορικά στοιχεία για τις προηγούμενες εκτελέσεις της τρισδιάστατης απεικόνιση. Επομένως το τμήμα σύστασης είναι σε θέση να προτείνει ποια πολιτική απέδωσε καλύτερα στο παρελθόν και ποιες από τις προηγούμενες αποφάσεις προκάλεσαν παραβιάσεις SLA.

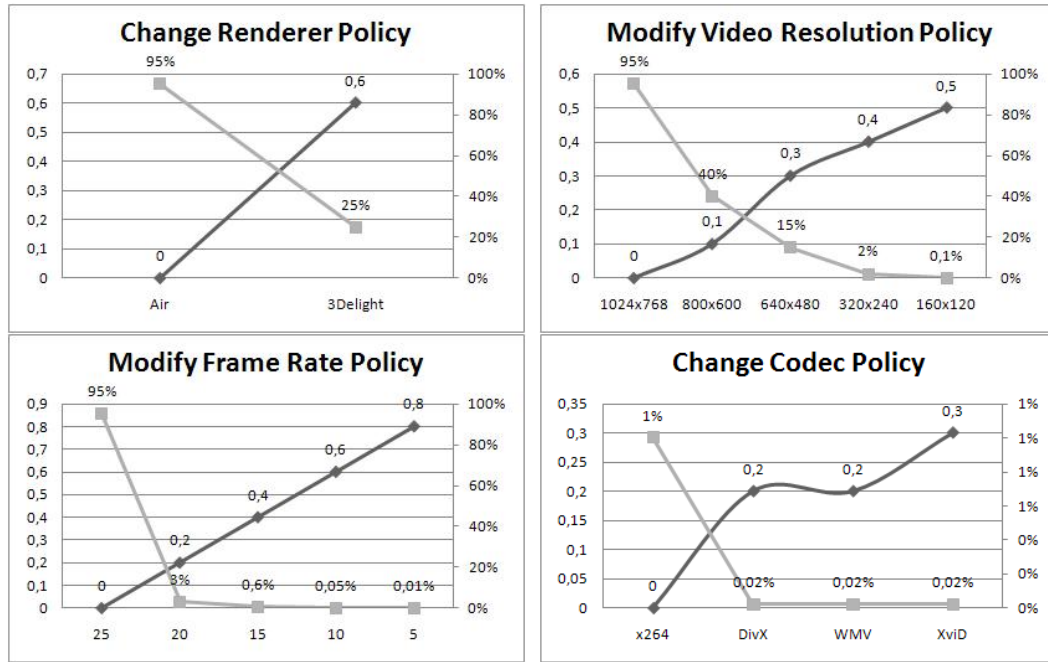
Η δοκιμή είναι βασισμένη στο GRIA 5 middleware [62] που προσαρμόστηκε αρκετά από την πλευρά φορέων παροχής υπηρεσιών προκειμένου να παραχθούν όλα τα απαραίτητα στοιχεία ελέγχου πραγματικού χρόνου στο αντίστοιχο component του framework παροχής QoS. Ο φορέας παροχής υπηρεσιών περιέλαβε οκτώ (8) υπολογιστικούς κόμβους συν έναν κόμβο αποθήκευσης δεδομένων, ο οποίος ήταν αρμόδιος για το video streaming. Προκειμένου να αξιολογηθεί η απόδοση του πλαισίου εργασίας παροχής QoS, αλλάχτηκαν τεχνητά οι παράμετροι χρόνου εκτέλεσης (μερικοί υπολογιστικοί κόμβοι ήταν ασύνδετοι). Το σύστημα θα έπρεπε αυτόματα και σε ένα σύντομο χρονικό πλαίσιο να επαναπροσδιοριστεί ώστε να εγγυηθεί το υψηλότερο πιθανό QoS και να ελαχιστοποιηθεί η πιθανότητα για την συνολική αποτυχία της εργασίας.

Αρχικά οι παράμετροι ελέγχου διαβιβάστηκαν στην Αξιολόγηση του Γεγονότος που αγνόησε μερικές από αυτές, όπως την αποθήκευση στοιχείων και τη διαθέσιμη μνήμη, και επισήμανε ως σημαντικότερες, την διαθεσιμότητα των υπολογιστικών κόμβων και το μήκος σειρών αναμονής των πλαισίων προς εκτέλεση. Στη συνέχεια, ο *επόπτης QoS* προσδιόρισε το σχετικό SLA και τις διαθέσιμες πολιτικές και ζήτησε τη σύσταση για το πώς αυτές οι πολιτικές εκτελέστηκαν στο παρελθόν και την πιθανότητα για μια παραβίαση SLA, το οποίο σε αυτήν την περίπτωση σημαίνει ότι το video streaming διακόπτεται. Το SLA περιλαμβάνει τις συγκεκριμένες αποζημιώσεις για τη συνολική αποτυχία της εργασίας (χάριν αυτού του σεναρίου η αξία αποζημιώσεων είναι "1 ευρώ

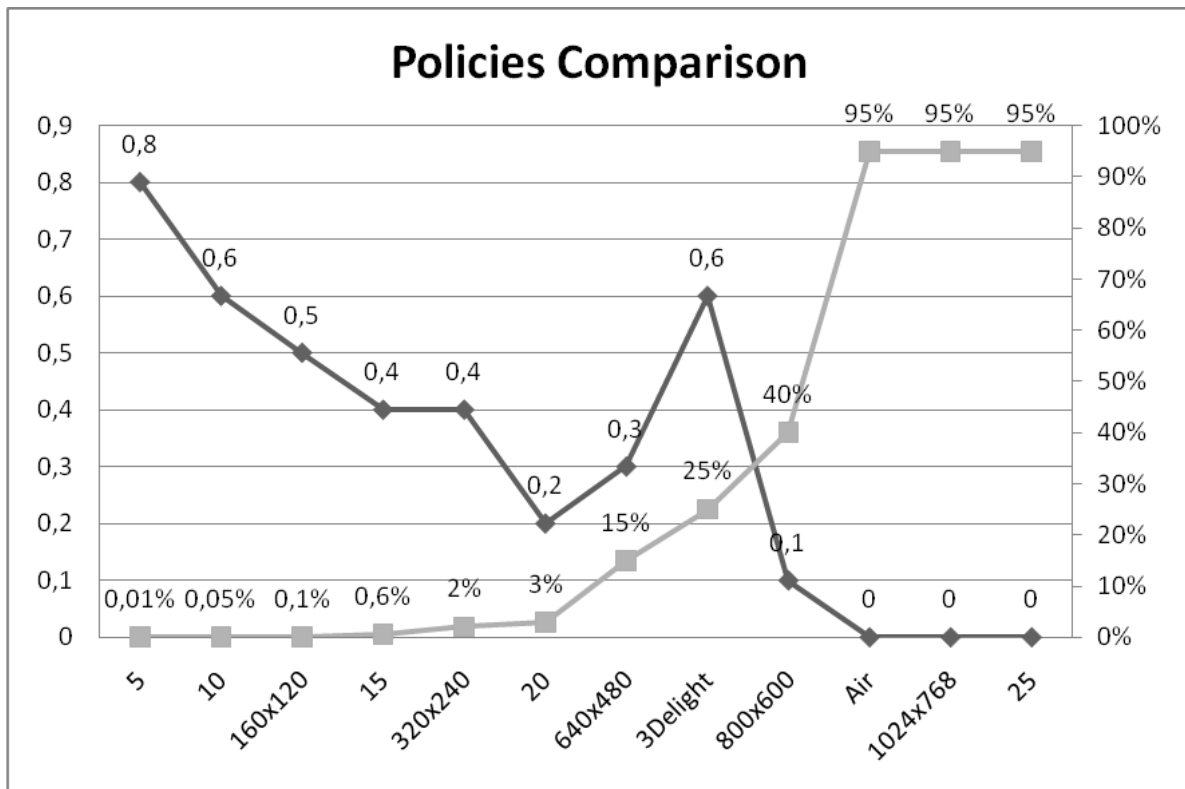
ανά δευτερόλεπτο"). Στους ακόλουθους αριθμούς απεικονίζουμε το συμπληρωματικό κόστος (τετραγωνική σειρά) που οφείλεται στην αλλαγή του επιπέδου QoS σε κάθε πολιτική και της κατ' εκτίμηση πιθανότητας για μια παραβίαση SLA (rhomb σειρά) βασισμένη στη σύσταση προηγούμενης εμπειρίας.

Όπως διευκρινίζεται σε αυτούς τους αριθμούς, αρχικά το συμπληρωματικό κόστος είναι μηδέν (αυτό συμφωνείται στο SLA) αλλά δεδομένου ότι η πιθανότητα παραβίασης SLA είναι υψηλή (95%) το σύστημα πρέπει να μετατραπεί εφαρμόζοντας μιας από τις πολιτικές παροχής QoS που συμπεριλαμβάνονται στο Σχήμα 27 και επίσης απεικονισμένες στο Σχήμα 28 (π.χ. Change Renderer Policy, Modify Video Resolution Policy, κ.λπ). Η ρύθμιση του επιπέδου QoS έχει επιπτώσεις και στο κόστος και στην πιθανότητα παραβίασης SLA και η καλύτερη επιλογή για το φορέα παροχής υπηρεσιών είναι να ελαχιστοποιήσει και τους δύο παράγοντες. Πρέπει να αναφερθεί ότι η πολιτική "Change Codec" δεν συσχετίζεται με την τρέχουσα κατάσταση του συστήματος και επομένως αποκλείεται από τις διαθέσιμες επιλογές.





Σχήμα 28: Επιδόσεις Εφαρμογής Πολιτικών



Σχήμα 29: Συνολική Επίδοση Εφαρμογής Πολιτικών

*Η σελίδα αυτή είναι σκόπιμα λευκή*

Το ανωτέρω διάγραμμα (Σχήμα 29) παρουσιάζει την πιθανότητα δαπανών και παραβίασης SLA για τις σχετικές πολιτικές και όπως μπορεί να φανεί η καλύτερη επιλογή θα ήταν να μειωθεί το ποσοστό πλαισίων σε 20 πλαίσια ανά δευτερόλεπτο δεδομένου ότι η πιθανότητα παραβίασης SLA υπολογίζεται σε 3% και το συμπληρωματικό κόστος για το φορέα παροχής υπηρεσιών λόγω της αλλαγής του επιπέδου QoS είναι 0,2 ευρώ ανά δευτερόλεπτο. Η παρούσα απόφαση μειώνει αρκετά την ποινική ρήτρα για το φορέα παροχής υπηρεσιών και επιπλέον επιτρέπει στο σύστημα για να μεταπηδήσει από αυτήν την κατάσταση στη συμφωνηθείσα, όπως δηλώνεται μέσα στο SLA.

## **6.5 Συμπεράσματα**

Σε αυτό το κεφάλαιο παρουσιάσαμε μια καινοτόμο προσέγγιση για την παροχή QoS σε SOA, ικανή να εγγυηθεί την υψηλή ζήτηση QoS των real-time επιχειρησιακών εφαρμογών χρησιμοποιώντας τις πληροφορίες ελέγχου σε πραγματικό χρόνο και των συστάσεων βασισμένων στην προηγούμενη εμπειρία. Περιγράφηκε επίσης μια υλοποίηση αναφοράς αυτής της προσέγγισης μαζί με ένα σενάριο επικύρωσης για μια πραγματική εφαρμογή (τρισεδιάστατη απεικόνιση) που καταδεικνύει και αξιολογεί τη λειτουργία της προτεινόμενης προσέγγισης.

Ο μηχανισμός παροχής QoS είναι σε θέση να αποκριθεί με το κατάλληλο, για τους τρέχοντες όρους, σύνολο ενεργειών προκειμένου να εγγυηθεί το πιο υψηλό επίπεδο QoS με μια λογική πιθανότητα παραβίασης SLA βασισμένη στις διαθέσιμες πολιτικές και μια σύσταση για το πώς το σύστημα έχει εκτελεσθεί στο παρελθόν. Η αξιολόγηση απέδωσε πολλά υποσχόμενα αποτελέσματα και επομένως η απόδοση αυτού του αυτόνομου και προσαρμοστικού πλαισίου θεωρούνται πλήρως εμπειριστατωμένα επιτρέποντας την υιοθέτηση του από οποιοδήποτε ετερογενές πραγματικού χρόνου και ειδικά περιβάλλον πλέγματος που επιδιώκει να φέρει τη γνώση QoS στο χρόνο εκτέλεσης.

Μέσα στα μελλοντικά σχέδιά μας να εφαρμοστεί αυτή η προσέγγιση στις ροές εργασιών εφαρμογής [19], από την οποία ανακύπτει ένα σύνολο ζητημάτων όπως οι συνδυασμοί διαθέσιμων πολιτικών ώστε να εγγυηθεί το end-to-end QoS. Ο μηχανισμός σύστασης θα μπορούσε επίσης να επεκταθεί με τις τεχνικές που εφαρμόστηκαν στα Συστήματα Ανιχνεύσεων Παρέισφρησης Ασφάλειας (Security Intrusion Detections Systems) δεδομένου ότι υπάρχουν πολλές ομοιότητες στο προτεινόμενο πλαίσιο παροχής QoS. Επιπλέον, τα αρχιτεκτονικά στοιχεία θα μπορούσαν επίσης να χρησιμοποιηθούν για την αξιολόγηση και συνεπώς το δυναμικό και σε πραγματικό χρόνο καθορισμό των διαθέσιμων πολιτικών για να υποστηρίξουν την αποδοτικότερη διαχείριση εργασίας μέσα στο φορέα παροχής υπηρεσιών από την άποψη του QoS.

# 7

## *Υπηρεσιοστρεφείς υποδομές για διαδραστικές εφαρμογές πραγματικού χρόνου*

Στο κεφάλαιο αυτό παρουσιάζουμε μια καινοτόμο υπηρεσιοστρεφή υποδομή ικανή να υποστηρίξει τις διαδραστικές εφαρμογές με απαιτήσεις απόκρισης σε πραγματικό χρόνο. Για τον αρχιτεκτονικό σχεδιασμό της υποδομής αναλύθηκαν οι απαιτήσεις των χρηστών εφαρμογών πραγματικού χρόνου για να προσδιοριστούν οι βασικές λειτουργίες και οι δυνατότητές της. Επιπλέον τα επιχειρηματικά μοντέλα της υποδομής καθώς και τα θέματα ασφάλειας εξετάστηκαν ώστε να σχεδιαστεί με τέτοιο τρόπο ώστε είναι οικονομικά βιώσιμη και φυσικά ασφαλής. Η πλατφόρμα αυτή σχεδιάστηκε και αναπτύχθηκε στα πλαίσια του ευρωπαϊκού ερευνητικού έργου IRMOS [51] ενώ η υλοποίησή της βρίσκεται σε εξέλιξη.

### **7.1 Αρχές σχεδιασμού της αρχιτεκτονικής**

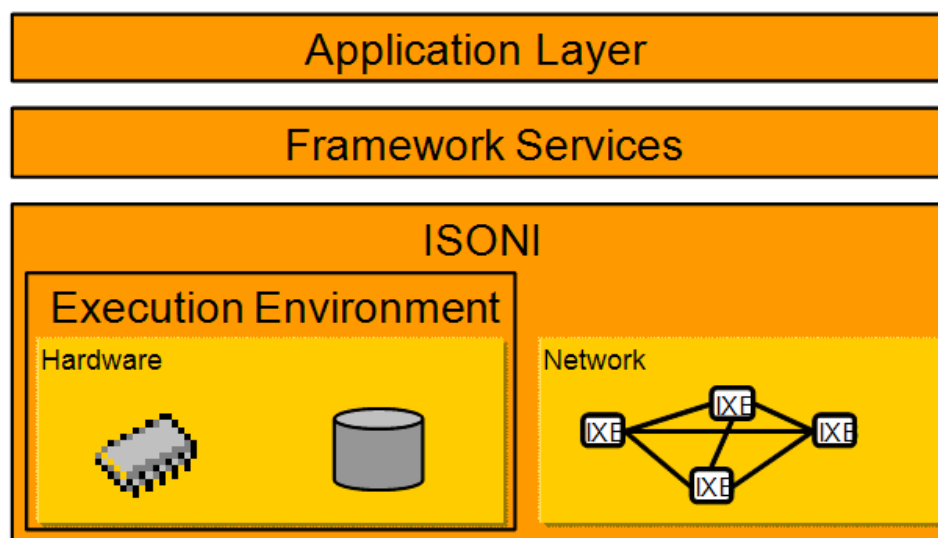
Ο κύριος στόχος του υποδομής είναι η προσαρμογή, «φιλοξενία» (hosting) και λειτουργία εφαρμογών πραγματικού χρόνου όπως η εικονική πραγματικότητα, τα πολυμέσα και το e-learning.

Οι υπηρεσιοστρεφείς υποδομές είναι σήμερα αναπόσπαστο κομμάτι των τεχνολογιών πληροφορικής και επικοινωνιών (ICT) και σίγουρα η τάση σχεδιασμού και υλοποίησης για τα συστήματα της επόμενης γενιάς. Επιπλέον, δεδομένου ότι η υποδομή στοχεύει να

διαχειριστεί και χρησιμοποιήσει διάφορους τύπους πόρων (δίκτυο, αποθήκευση, υπολογιστικούς) μέσα από τεχνολογίες virtualization [56] καθώς επίσης και να επικοινωνεί σε πραγματικό χρόνο με συστήματα άλλων τεχνολογιών όπως οι Wifi locators [74], η υπηρεσιοστρεφής αρχιτεκτονική προσέγγιση αποτελεί τη μόνη επιλογή για την υποδομή. Από τη μία, οι υποδομές SOI έχουν διάφορα πλεονεκτήματα, όπως θα δούμε πιο κάτω, προσφέροντας υπηρεσίες που αλληλεπιδρούν δυναμικά και συνεχώς, εκτείνονται μεταξύ των διαφορετικών περιοχών διαχείρισης (domains) από το επίπεδο της εφαρμογής μέχρι το επίπεδο διαχείρισης των πόρων, του δικτύων και του περιβάλλοντος λειτουργίας. Επομένως, η πρόκληση είναι να σχεδιαστεί προσεκτικά και να συγχρονιστεί αυτό το πλούσιο σύνολο υπηρεσιών ώστε να ενεργοποιηθούν αποτελεσματικά και να διαχειριστούν όλοι οι πόροι σε πραγματικό χρόνο παρέχοντας στους τελικούς χρήστες και τις εφαρμογές τους την ποιότητα υπηρεσιών που έχει συμφωνηθεί στα SLAs. Σε αυτό το πλαίσιο, ένα από τα βασικά ζητούμενα είναι η αυτοματοποιημένη και δυναμική διαπραγμάτευση των SLAs τόσο στο επίπεδο της εφαρμογής όσο και στο επίπεδο των πόρων ώστε να διασφαλιστεί η ποιότητα υπηρεσιών σε κάθε σημείο της αλυσίδας προστιθέμενης αξίας (value chain).

Οι υπηρεσίες της υποδομής εκτείνονται σε διάφορα επίπεδα[84]. Έτσι υπάρχουν υπηρεσίες που υποστηρίζουν τους υπεύθυνους για την ανάπτυξη εφαρμογής (application developers) στην μετατροπή των εφαρμογών τους στο περιβάλλον εκτέλεσης της υποδομής ενώ άλλες υπηρεσίες θα υποστηρίζουν, σε πραγματικό χρόνο, την εκτέλεση εφαρμογών. Επομένως, η υποδομή είναι δεν είναι μόνο ένα σύνολο υπηρεσιών που αλληλεπιδρούν μεταξύ τους, όπως στα κοινά SOIs, αλλά ένα real-time SOI στο οποίο οι υπηρεσίες με δυνατότητες πραγματικού χρόνου επεκτείνονται στα εικονικά περιβάλλοντα (virtual environments) που εποπτεύουν τον κύκλο της ζωής εφαρμογής και διασφαλίζουν το συμφωνηθέν επίπεδο QoS.

Η συνολική διαχείριση και ο έλεγχος της υποδομής και των μεμονωμένων υπηρεσιών έχουν ως σκοπό να υποστηρίξουν σε πραγματικό χρόνο την αλληλεπίδρασή τους, συμπεριλαμβανομένης επίσης της αλληλεπίδρασης μεταξύ των εφαρμογών και των πόρων. Η συγκεκριμένη υποδομή αναμένεται για να δώσει λύση σε διάφορα προβλήματα σχετικά με τη διασφάλιση της ποιότητας υπηρεσιών, το σχεδιασμό και δρομολόγηση (scheduling) της εκτέλεσης των υπηρεσιών, και την αντιστοίχιση των ροών εργασιών και των απαιτήσεων της εφαρμογής σε χαμηλού επιπέδου παραμέτρους των πόρων και μεμονωμένες υπηρεσίες. Τα προαναφερθέντα ζητήματα θα αντιμετωπιστούν τόσο στο επίπεδο των υπηρεσιών πλαισίου όσο και στο επίπεδο του εικονικού περιβάλλοντος (virtual environment) και εγγυώνται απόδοση πραγματικού χρόνου συνολικά στην πλατφόρμα. Να σημειώσουμε ότι στα πλαίσια του έργου IRMOS το εικονικό περιβάλλον που χρησιμοποιείται ονομάζεται ISONI (Intelligent Service Oriented Network Infrastructure) [52]. Το ακόλουθο διάγραμμα απεικονίζει τη γενική αρχιτεκτονική της υποδομής:



*Σχήμα 30: Υπηρεσιοστρεφής υποδομή πραγματικού χρόνου*

*Η σελίδα αυτή είναι σκόπιμα λευκή*



### **7.1.1 Υπηρεσιοστρεφείς υποδομή**

Οι υπηρεσιοστρεφείς υποδομές βασίζονται στον ανασχηματισμό των λειτουργιών και των πληροφοριών μιας εφαρμογής σε μια «υπηρεσία» η οποία μπορεί να προσεγγιστεί μέσω μιας κοινής διεπαφής ανεξάρτητα από τη θέση, τη λειτουργία, ή των δεδομένων επεξεργασίας.

Τα γενικά χαρακτηριστικά της υποδομής είναι [70] :

- **Χαλαρά διασυνδεδεμένες υπηρεσίες:** Η χαλαρή διασύνδεση είναι μια προσέγγιση σχεδιασμού κατανεμημένων εφαρμογών προσφέροντας ευκινησία, τη δυνατότητα δηλαδή της υποδομής να προσαρμοστεί εύκολα και γρήγορα στις αλλαγές. Η χαλαρή διασύνδεση θυσιάζει σκόπιμα κάποιες φορές τη βελτιστοποίηση των διεπαφών για να επιτύχει εύκαμπτη διαλειτουργικότητα μεταξύ συστημάτων που είναι ανόμοια σε τεχνολογία, θέση, απόδοση και διαθεσιμότητα. Με αυτό τον τρόπο μια αόριστα διασυνδεδεμένη εφαρμογή απομονώνει τις εσωτερικές της αλλαγές σε μέσω της διεπαφής της προς τις άλλες εφαρμογές. Σε σύγκριση με τις παραδοσιακές, στενά συνδεδεμένες εφαρμογές, οι αόριστα συνδεδεμένες εφαρμογές στοχεύουν στο να είναι περισσότερο επαναχρησιμοποιήσιμες και προσαρμόσιμες.
- **Σύγχρονες:** Με τη χρήση σύγχρονου τρόπου επικοινωνίας μεταξύ των υπηρεσιών όταν ο χρήστης ζητά πληροφορίες ή επικαλείται μια λειτουργία, η διασύνδεση διατηρείται έως το τέλος της διαδικασίας.
- **Ασύγχρονες:** Με τον ασύγχρονο τρόπο επικοινωνίας οι υπηρεσίες στέλνουν τις οποίες πληροφορίες χωρίς να αναμένεται να πάρουν μια πίσω άμεση απάντηση. Αυτό το χαρακτηριστικό είναι πολύ σημαντικό στις περιπτώσεις όπου δεν απαιτείται να διατηρηθεί μια σύνδεση μεταξύ δύο συστημάτων περιμένοντας μια απάντηση.

Η σημαντικότερη πτυχή στην προτεινόμενη αρχιτεκτονική είναι ότι ο σχεδιασμός της πραγματοποιήθηκε βάσει των real-time απαιτήσεων των εφαρμογών. Κατά συνέπεια, οι υπηρεσίες και η σύνθεσή τους έγινε με τέτοιο τρόπο ώστε να διατηρηθούν συνολικά οι real-time ιδιότητες της υποδομής συμπεριλαμβανομένων των πόρων, των εικονικών περιβαλλόντων εκτέλεσης καθώς επίσης και των δικτύων επικοινωνιών.

### **7.1.2 Απόκριση σε πραγματικό χρόνο**

Σε ένα υπολογιστικό σύστημα πραγματικού χρόνου, η ποιότητα του αποτελέσματος εξαρτάται όχι μόνο από την ακρίβεια της παραγωγής, αλλά και στο χρόνο ότι αυτό παράγεται. Σε ένα hard real-time σύστημα, ακόμη και η μικρότερη παραβίαση μπορεί να είναι καταστρεπτική για τη γενική λειτουργία του συστήματος. Από την άλλη, σε ένα soft real-time σύστημα, μερικές παραβιάσεις των χρονικών περιορισμών είναι αποδεκτές, εφ' όσον αυτές δεν είναι πάρα πολύ συχνές και παραμένουν μέσα στα αποδεκτά όρια. Η συγκεκριμένη υποδομή είναι ένα τέτοιο soft real-time σύστημα πράγμα που σημαίνει ότι πρέπει να ενσωματώσει τις λειτουργίες πραγματικού χρόνου τα σε όλα στρώματά του. Γενικά όλες οι αλληλεπιδράσεις μεταξύ των διαφόρων συστατικών της υποδομής θα πρέπει να είναι οι ελάχιστες δυνατές ώστε να ελαχιστοποιηθούν κατά αυτόν τον τρόπο και οι λανθάνουσες καταστάσεις κατά τη λειτουργία της υποδομής.

Ένα σημαντικό μέρος της υποδομής είναι το εικονικό περιβάλλον εκτέλεσης που είναι η βάση για την παροχή της σε πραγματικό χρόνο λειτουργίας. Σε αυτό το επίπεδο, οι υπολογιστικοί κόμβοι πάνω στους οποίους χτίζεται το εικονικό περιβάλλον εκτέλεσης περιλαμβάνει αρκετά χαρακτηριστικά λειτουργίας σε πραγματικό χρόνο όπως real-time linux kernel, real-time scheduling, live migration των virtual environments κ.α. Παρά το γεγονός ότι το χαμηλό επίπεδο είναι «real-time enabled» ανεπιθύμητες καταστάσεις είναι πιθανό να εμφανιστούν, απαιτούνται ισχυροί μηχανισμοί ανίχνευσης λαθών και αποκατάστασης της ομαλής λειτουργίας. Αυτό μπορεί να έχει μεγάλη επίδραση στις real-

time ικανότητες της πλατφόρμας, δεδομένου ότι οι ευφυείς μηχανισμοί αποκατάστασης λαθών θα επιτρέψουν στην εφαρμογή να παρέχει την επιθυμητή ποιότητα υπηρεσιών και να ικανοποιήσει τους χρονικούς περιορισμούς ακόμα και σε περίπτωση προσωρινής αποτυχίας.

Πριν το σχεδιασμό της αρχιτεκτονικής της υποδομής προσδιορίσαμε και αναλύσαμε τα σημεία της υποδομής που είναι απαραίτητα η λειτουργία πραγματικού χρόνου. Γενικά, η λειτουργία πραγματικού χρόνου απαιτείται για τις διαδικασίες εκτέλεσης και ελέγχου για τα δίκτυα υπηρεσιών, ενώ για τις διαδικασίες προ-εκτέλεσης όπως η διαχείριση SLA, η λειτουργία πραγματικού χρόνου δεν είναι κρίσιμη.

Από την ανάλυση των απαιτήσεων των χρηστών και των περιπτώσεων χρήσης (use cases) καταλήξαμε στο συμπέρασμα είναι ότι η λειτουργία πραγματικού χρόνου και απαιτείται στα υποσυστήματα όπως η ευφυής δικτύωση, οι υποδομές αποθήκευσης, οι κόμβοι εκτέλεσης εργασιών, και στις υπηρεσίες πλαισίου που υποστηρίζουν την εκτέλεση των εφαρμογών. Με αυτόν τον τρόπο, η λειτουργία πραγματικού χρόνου εξασφαλίζεται από το επίπεδο των πόρων (π.χ. πυρήνας Linux, δρομολογητές δικτύων) μέχρι το επίπεδο των υπηρεσιών από το εικονικό περιβάλλον μέχρι την εφαρμογή. Επιπλέον, κάθε υποσύστημα θα μπορεί να ρυθμίζεται αυτόνομα προκειμένου να αποφευχθεί η κλιμάκωση ενός προβλήματος (όπως μια αποτυχία συνδέσεων παραβίασης SLA ή δικτύων) σε ανώτερο επίπεδο λόγω του ότι κάτι τέτοιο μπορεί να προκαλέσει καθυστερήσεις. Από την άλλη πλευρά όμως αυτό απαιτεί πρόσθετη λειτουργικότητα για μερικά υποσυστήματα ή δομικά συστατικά όπως το σύστημα αποθήκευσης και το εικονικό περιβάλλον. Επίσης οι διαδικασίες που περιλαμβάνουν περισσότερα από ένα δομικά συστατικά, πρέπει να καθοριστούν με τρόπο που να μην εμποδίζει την λειτουργία πραγματικού χρόνου της υποδομής. Τέλος, οι πλατφόρμες υλικού και το λογισμικό στις οποίες θα λειτουργήσουν οι εφαρμογές θα πρέπει πληρούν προδιαγραφές λειτουργίας πραγματικού χρόνου.

## **7.2 Συμφωνίες επιπέδου υπηρεσιών**

### **7.2.1 Διαχείριση συμφωνιών επιπέδου υπηρεσιών**

Η διαχείριση των συμφωνιών επιπέδου υπηρεσιών (SLA) καλύπτει διαφορετικές φάσεις λειτουργίας της υποδομής: α) τη δημοσίευση περιγραφών των υπηρεσιών και των δυνατοτήτων τους, τη διαπραγμάτευση και τη καθιέρωση των συμφωνιών, την παροχή και επίβλεψη των υπηρεσιών και των πόρων και τέλος την αξιολόγηση και επιβολή των όρων που περιλαμβάνουν τα SLAs.

Η διαχείριση μιας συμφωνίας επιπέδων υπηρεσιών είναι ένας σύνθετος στόχος, όπως διασφάλιση παροχής ποιότητα υπηρεσιών, και εξαρτάται από διάφορες πτυχές όπως η συμπεριφορά του συστήματος, η αξιοπιστία δικτύων, οι εξωτερικές εξαρτήσεις και τέλος τυχόν απροσδόκητα γεγονότα που μπορεί να συμβούν. Αυτό κάνει τη διαχείριση SLAs έναν σύνθετο στόχο που έχει επιπτώσεις στον τρόπο διαχείρισης των πόρων της υποδομής. Σε περίπτωση διακοπής λειτουργίας ή αποτυχίας των υπηρεσιών που θα προκαλέσει αναπόφευκτα μια παραβίαση SLA, ο φορέας παροχής υπηρεσιών πρέπει να πάρει τις αποφάσεις βασισμένες στις επιχειρηματικές σχέσεις του με τους πελάτες, την τρέχουσα κατάσταση υποδομής και τις ενδεχόμενες ποινικές ρήτρες, προκειμένου να ελαχιστοποιηθεί το συνολικό κόστος που προκύπτει από την παραβίαση του SLA.

Η διαπραγμάτευση SLA αναφέρεται στα διαφορετικά στάδια που οδηγούν τους πελάτες και τους προμηθευτές σε συμφωνία σχετικά με ένα SLA. Η διαδικασία αρχίζει με τον πελάτη που διευκρινίζει τους επιθυμητούς όρους απόδοσης και συνεχίζει με τους μηχανισμούς αναζήτησης υπηρεσιών και τις μεθόδους (επανα-)διαπραγμάτευσης υπηρεσιών με τη χρήση συγκεκριμένων πολιτικών. Ο πελάτης και ο προμηθευτής έχουν μια κοινή γλώσσα έκφρασης μέσω της χρήσης συγκεκριμένων προτύπων σύνταξης των συμφωνιών (SLA templates). Μόλις υπάρξει μια συμφωνία το SLA θεωρείται ενεργό, περιέχοντας τις πληροφορίες για το συμφωνηθέν επίπεδο QoS και την τιμή. Από την

πλευρά της υποδομής απαιτείται συνεχής έλεγχος και προσαρμογή των διαδικασιών προκειμένου ο πάροχος να εγγυηθεί το επίπεδο QoS που απαιτείται από το χρήστη.

### **7.2.2 Τύποι συμφωνιών επιπέδου υπηρεσιών και η σχέση τους με την αλυσίδα προστιθέμενης αξίας**

Στο πλαίσιο της υποδομής (IRMOS) υπάρχουν διαφορετικοί δράστες που τοποθετούνται σε διαφορετικά επίπεδα: έχουμε τους καταναλωτές/πελάτες, τους παρόχους εφαρμογής, τους παρόχους της υπηρεσιοστρεφούς υποδομής και τους παρόχους του εικονικού περιβάλλοντος. Να σημειώσουμε ότι η σχέση μεταξύ των δραστών είναι «πολλά προς πολλά» (N:N) που σημαίνει ότι για κάθε πάροχο μπορεί να έχουμε περισσότερους του ενός χρήστες αλλά και το αντίστροφο. Η αλυσίδα προστιθέμενης αξίας (value chain), βάση της οποίας αναπτύχθηκε η υποδομή, παρουσιάζεται στο Σχήμα 31 ενώ οι δράστες και ο ρόλος του κάθε ενός στη αλυσίδα προστιθέμενης αξίας περιγράφεται λεπτομερώς παρακάτω:

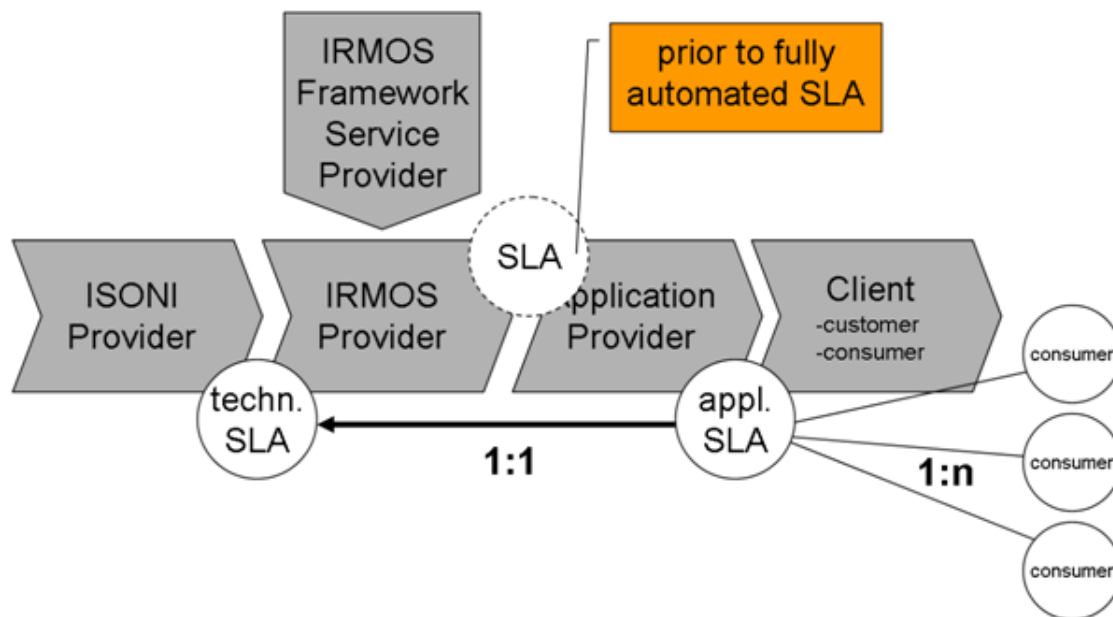
- Ο πάροχος εικονικού περιβάλλοντος (ISONI Provider) λειτουργεί και διαχειρίζεται την υποδομή εικονικού περιβάλλοντος και τους πόρους που αυτή περιλαμβάνει (υπολογιστικούς, δίκτυο και αποθηκευτικούς). Η πάροχος εικονικού περιβάλλοντος δεν έχει καμία απολύτως επίγνωση της εφαρμογής που εκτελείται.
- Ο πάροχος υπηρεσιοστρεφούς υποδομής (IRMOS Provider) λειτουργεί και διαχειρίζεται τις υπηρεσίες πλαισίου (framework services). Οι υπηρεσίες πλαισίου είναι το ενδιάμεσο επίπεδο μεταξύ της εφαρμογής και των πόρων που αυτές εκτελούνται ενώ επιπλέον παρέχει όλους τους μηχανισμούς που απαιτούνται από τους προμηθευτές της κάθε εφαρμογής προκειμένου να την προσαρμόσουν στην υποδομή.

- Ο πάροχος λογισμικού υπηρεσιών πλαισίου (IRMOS Framework Service Provider) αναπτύσσει και παρέχει το λογισμικό των υπηρεσιών πλαισίου στους παρόχους υπηρεσιοστρεφούς υποδομής.
- Ο πάροχος εφαρμογής (Application Provider) αλληλεπιδρά με τις υπηρεσίες πλαισίου ώστε να προσφέρει την εφαρμογή μέσω της υπηρεσιοστρεφούς υποδομής στους τελικούς χρήστες.
- Ο πελάτης (Client) χρησιμοποιεί τις διεπαφές που παρέχονται από τον πάροχο εφαρμογής για να αλληλεπιδράσουν με την εφαρμογή, τον καθορισμό μιας των SLA κ.α.

Οι διαφορετικές απαιτήσεις αυτών των πέντε διαφορετικών δραστών κατά τη διάρκεια της εκτέλεσης μιας εφαρμογής στην υποδομή δεν μπορούν να συνοψιστούν σε ένα ενιαίο SLA. Επιπρόσθετα οι παράμετροι που περιλαμβάνονται στα SLA δεν είναι οι ίδιες για τον πελάτη και τον πάροχο εικονικού περιβάλλοντος. Αυτός είναι και ο κύριος λόγος που απαιτούνται SLA μεταξύ των δραστών σε διάφορα επίπεδα της υποδομής. Συγκεκριμένα προσδιορίσαμε τους παρακάτω τύπους SLAs:

- **Συμφωνία εφαρμογής (Application SLA / A-SLA):** Πρόκειται για μία συμφωνία που συνάπτεται μεταξύ του πελάτη και του παρόχου εφαρμογής. Αυτοί οι δράστες αντιπροσωπεύουν τον καταναλωτή/πελάτη των εφαρμογών που πρόκειται να επεκταθούν και να λειτουργήσουν στην υποδομή καθώς και τον πάροχο που προσφέρει την εφαρμογή μέσω της υποδομής. Αυτό το SLA περιέχει τις υψηλού επιπέδου παραμέτρους ποιότητας υπηρεσιών της εφαρμογής όπως αυτές καθορίζονται από το χρήστη.
- **Τεχνική συμφωνία (Technical SLA / T-SLA):** Αυτή η συμφωνία είναι μεταξύ του παρόχου εικονικού περιβάλλοντος και του παρόχου υπηρεσιοστρεφούς

υποδομής. Αντιλαμβανόμαστε ότι αυτή η συμφωνία περιέχει τις χαμηλού επιπέδου παραμέτρους ποιότητας υπηρεσιών που συνδέονται με την υποδομή. Η μετάφραση των υψηλού επιπέδου παραμέτρων που προέρχονται από της συμφωνίες εφαρμογής στις χαμηλού επιπέδου παραμέτρους πραγματοποιείται από τις υπηρεσίες πλαισίου και συγκεκριμένα από τους μηχανισμούς αντιστοίχισης ποιότητας υπηρεσιών.



Σχήμα 31: Οι τύποι συμφωνιών επιπέδου υπηρεσιών

Να σημειώσουμε ότι υπάρχει ένα επιπλέον SLA μεταξύ του παρόχου της εφαρμογής και του παρόχου υπηρεσιοστρεφούς υποδομής. Πρόκειται για μία μακροπρόθεσμη σχέση που δημιουργείται όταν μία εφαρμογή εντάσσεται στην υποδομή ενώ προηγείται της αυτοματοποιημένης διαδικασία διαπραγμάτευσης συμφωνιών εφαρμογής και τεχνικών συμφωνιών.

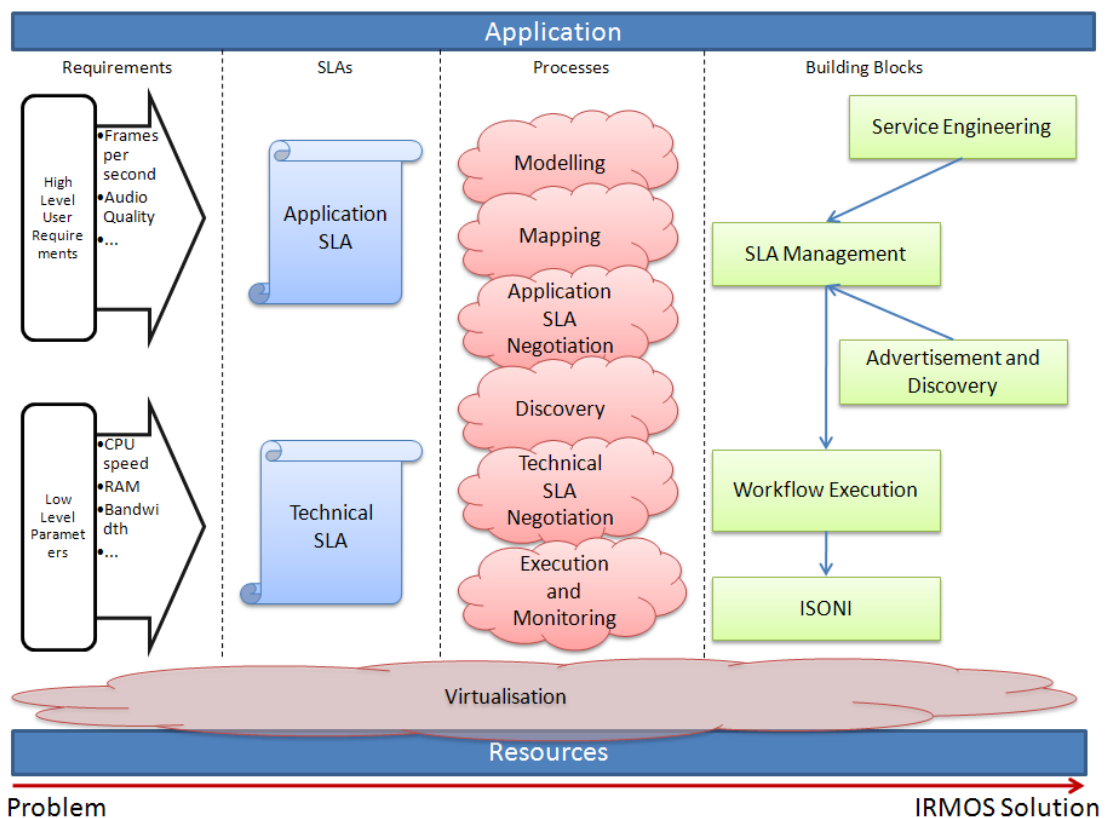
### 7.2.3 Καινοτομίες στη διαχείριση συμφωνιών επιπέδου υπηρεσιών

Η πραγματικού χρόνου υπηρεσιοστρεφής υποδομή υποστηρίζει τη λειτουργία εφαρμογών σε εικονικά περιβάλλοντα εποπτεύοντας τον κύκλο της ζωής των εφαρμογών αυτών ενώ

*Η σελίδα αυτή είναι σκόπιμα λευκή*



παράλληλα μπορεί να εγγυηθεί το συμφωνηθέν επίπεδο ποιότητας υπηρεσιών. Για να επιτευχθούν αυτά απαιτούνται συγκεκριμένες διαδικασίες οι οποίες σχετίζονται άμεσα και υλοποιούνται με συγκεκριμένα δομικά συστατικά της υποδομής. Η υποδομή παρέχει υπηρεσίες που υποστηρίζουν τους υπεύθυνους για την ανάπτυξη εφαρμογής στην προσαρμογή και μεταφορά την εφαρμογών στην υποδομή. Αυτό επιτρέπει τον καθορισμό συμφωνιών εφαρμογής (A-SLA) οι οποίες μπορούν να συσχετιστούν με τις χαμηλού επιπέδου παραμέτρους και στη συνέχεια να χρησιμοποιηθούν για την αναζήτηση και δέσμευση των διαθέσιμων πόρων και υπηρεσιών. Η δέσμευση αυτή παίρνει τη μορφή τεχνικής συμφωνίας (T-SLA) που θα χρησιμοποιηθεί κατά τη διάρκεια της εκτέλεσης της εφαρμογής για να διαμορφωθούν οι χαμηλού επιπέδου μηχανισμοί στο επίπεδο διαχείρισης πόρων των υπολογιστικών συστημάτων και της υποδομής δικτύου προκειμένου να παρασχεθούν τα επιθυμητά επίπεδα QoS και να ελεγχθεί η εκτέλεση των εφαρμογών. Όλα τα παραπάνω γίνονται με ένα πλήρως αυτοματοποιημένο τρόπο.



Σχήμα 32: Τα πλεονεκτήματα της υποδομής

*Η σελίδα αυτή είναι σκόπιμα λευκή*

Το προηγούμενο σχήμα παρουσιάζει τις διαδικασίες που ακολουθήσαμε στην υποδομή από την εφαρμογή μέχρι τους πόρους και από τις απαιτήσεις των χρηστών μέχρι τα δομικά συστατικά. Μέχρι τώρα, οι περισσότεροι μηχανισμοί SLA θεωρούν τη διαπραγμάτευση μια διαδικασία που πραγματοποιείται μόνο μία φορά μέσα στον κύκλο της ζωής του SLA, πριν από τη φάση εκτέλεσης. Μόλις ολοκληρωθεί η διαπραγμάτευση, οι υπηρεσίες και οι πόροι ελέγχονται αναφορικά με το SLA ενώ σε περίπτωση που υπάρχει παραβίαση σε κάποιο όρο, διάφορες ενέργειες (που προβλέπονται από το ίδιο το SLA στις περισσότερες περιπτώσεις) εκτελούνται. Εντούτοις, μερικές φορές οι αιτίες των παραβιάσεων θα μπορούσαν να αντιμετωπιστούν με επανάληψη της διαδικασίας της διαπραγμάτευσης. Αυτό ονομάζεται επαναδιαπραγμάτευση SLA ή διαπραγμάτευση κατά το χρόνο εκτέλεσης. Αυτό αποτελεί μια από τις καινοτόμες πτυχές διαχείρισης της υποδομής, γνωστό ως δυναμική συμφωνία επιπέδου υπηρεσιών (Dynamic SLA).

Μια άλλη σημαντική πτυχή στη διαχείριση SLA στην υποδομή σχετίζεται με τη φύση της διαδικασίας διαπραγμάτευσης. Αυτή η καινοτομία είναι η αυτόματη διαπραγμάτευση SLA, που σημαίνει ότι το κομμάτι της υποδομής που συμμετέχει στη διαδικασία διαπραγμάτευσης έχει σε κάποιο βαθμό την ικανότητα να αποφασίσει, χωρίς ανθρώπινη επέμβαση, για τη διαπραγμάτευση SLA. Η διαδικασία έχει βασιστεί σε πολιτικές που καθορίζονται από τους αντίστοιχους δράστες που συμμετέχουν στη διαπραγμάτευση.

Για να ολοκληρώσει όμως αυτόν τον στόχο η υπηρεσία διαχείρισης SLA στηρίζεται σε άλλες υπηρεσίες της υποδομής. Πολύ σημαντικές υπηρεσίες για τη διαχείριση SLA είναι η αντιστοίχιση παραμέτρων (Mapping) και η μέτρηση επιδόσεων (Benchmarking). Με την χρήση της υπηρεσία μέτρησης επιδόσεων παράγονται τα μοντέλα συμπεριφοράς της εφαρμογής που περιγράφουν τις ανάγκες της εφαρμογής για πόρους ανάλογα με τον τρόπο που χρησιμοποιείται (παράμετροι λειτουργίας, δεδομένα εισόδου, κτλ.). Αυτά τα μοντέλα χρησιμοποιούνται στη συνέχεια από την υπηρεσία αντιστοίχισης παραμέτρων, η

οποία είναι αρμόδια για τη μετάφραση των υψηλού επιπέδου απαιτήσεων, που καθορίζονται από το χρήστη στις συμφωνίες εφαρμογής, στις χαμηλού επιπέδου παραμέτρους που χρησιμοποιούνται στις τεχνικές συμφωνίες. Η σημασία της διαδικασίας αυτής συνοψίζεται στο ότι ο χρήστης δεν χρειάζεται να συνάψει ένα SLA που περιέχει ενδεχομένως παραμέτρους για τις οποίες δεν μπορεί να έχει την πλήρη κατανόηση (όπως ο κύκλος KME ή το εύρος ζώνης δικτύων), αλλά είναι σε θέση να καθορίζει τις απαιτήσεις του στο επίπεδο της εφαρμογής, κάτι με το οποίο είναι απόλυτα εξοικειωμένος (όπως τα καρτέ ανά δευτερόλεπτο ή την ανάλυση εικόνας). Από την άλλη πλευρά, ο προμηθευτής μπορεί να βεβαιωθεί ότι θα είναι σε θέση να ικανοποιήσει τις απαιτήσεις που τίθενται από το χρήστη για την ομαλή εκτέλεση της εφαρμογής.

#### ***7.2.4 Εικονοποίηση των πόρων και διασφάλιση της ποιότητας υπηρεσιών***

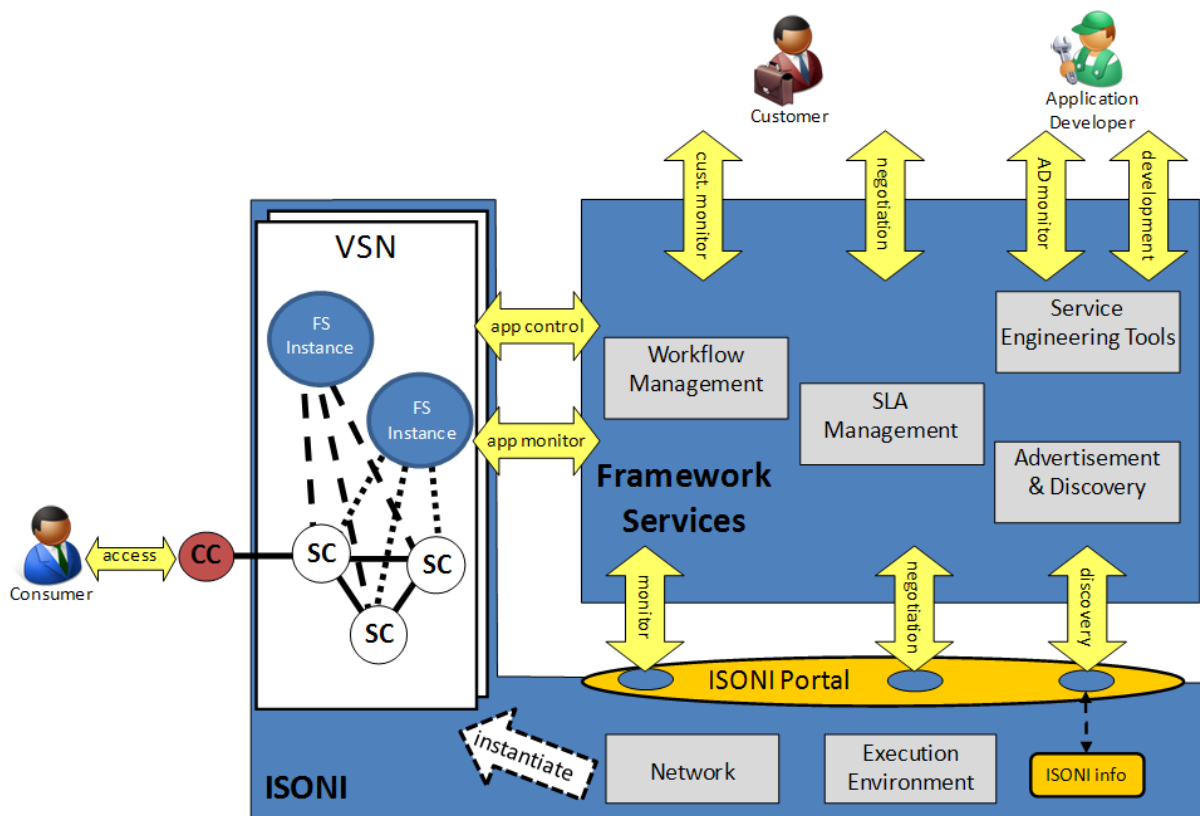
Πολύ σημαντικό στοιχείο της υποδομής είναι διαχείριση των πόρων από την πλευρά των παρόχων με τη χρήση τεχνικών εικονοποίησης (virtualization). Αυτό σημαίνει ότι, με τη χρήση εικονοποίησης, ο πάροχος των πόρων είναι σε θέση να προσφέρει ολοκληρωμένα περιβάλλοντα λειτουργίας των εφαρμογών χωρίς να χρειάζεται να γνωρίζει λεπτομέρειες σχετικά με το περιβάλλον αυτό (όπως π.χ. λειτουργικό σύστημα) ή για τις ίδιες τις εφαρμογές. Οι μόνες πληροφορίες που είναι σχετικές για τον πάροχο των πόρων είναι: ο τύπος πόρου που απαιτείται (π.χ. υπολογισμός, αποθήκευση), η τοπολογία των διασυνδέσεών τους και τέλος οι απαιτήσεις ποιότητας υπηρεσιών για τους πόρους και τις διασυνδέσεις τους.

Επιπρόσθετα με αυτό τον τρόπο ο πάροχος πόρων μπορεί να μοιραστεί τον ίδιο φυσικό πόρο (δηλ., κόμβος υπολογισμού) για τη φιλοξενία πολλαπλών εικονικών πόρων (δηλ., λειτουργικά συστήματα). Ενώ αυτό είναι ιδιαίτερος χρήσιμο από επιχειρηματικής πλευράς, προκύπτουν τεχνικά ζητήματα διαχείρισης των κόμβων προκειμένου να δοθούν οι κατάλληλες εγγυήσεις παροχής ποιότητας υπηρεσιών σε κάθε μεμονωμένο εικονικό

περιβάλλον. Η λύση σε αυτό το ζήτημα στηρίζεται στη χρήση κατάλληλων μηχανισμών δρομολόγησης πόρων, βελτιστοποιώντας, με τη χρήση πολιτικών την διαχείριση τους.

### 7.3 Καινοτόμος ορισμός της αρχιτεκτονικής

Ο στόχος της υποδομής είναι να εξασφαλίσει την ομαλή εκτέλεση διαδραστικών εφαρμογών πραγματικού χρόνου. Προηγούμενα παρουσιάσαμε τις αρχιτεκτονικές προδιαγραφές για την υποδομή καθώς και τα βασικά χαρακτηριστικά της. Από την ανάλυση αυτή προέκυψαν συγκεκριμένα υποσυστήματα για την υποδομή κάθε ένα από τα οποία αντιμετωπίζει διαφορετικά ζητήματα. Η γενική αρχιτεκτονική της υποδομής παρουσιάζεται στο ακόλουθο σχήμα:



Σχήμα 33: Η γενική αρχιτεκτονική

Στο σχήμα αυτό παρατηρούμε τα δυο βασικά υποσυστήματα της υποδομής: τις υπηρεσίες πλαισίου (Framework Services) και το εικονικό περιβάλλον (ISONI). Στο σχεδιασμό της

*Η σελίδα αυτή είναι σκόπιμα λευκή*

υποδομής ακολουθήσαμε μια καινοτόμο προσέγγιση αναφορικά με τον τρόπο που τα συστήματα αλληλεπιδρούν που είναι αρκετά διαφορετική από τα συμβατικές υπηρεσιοστρεφείς υποδομές SOA ή υποδομές πλέγματος κυρίως λόγω λειτουργίας σε πραγματικού χρόνου και των τεχνολογιών εικονοποίησης των πόρων. Σε πρώτη φάση, οι υπηρεσίες πλαισίου παρέχουν μηχανισμούς στον υπεύθυνο ανάπτυξης της εφαρμογής για την προσαρμογή της εφαρμογής όπως και μηχανισμούς διασφάλισης ποιότητας υπηρεσιών και διαπραγμάτευσης συμφωνιών επιπέδου υπηρεσιών για τον πάροχο της υποδομής. Οι χρήστες μπορούν να έχουν πρόσβαση στην υποδομή με δύο τρόπους:

- Ο πελάτης μπορεί να διαπραγματεύεται συμφωνίες εφαρμογής και ελέγχει τον κύκλο της ζωής εφαρμογής μέσω των υπηρεσιών πλαισίου.
- Ο καταναλωτής μπορεί να έχει πρόσβαση και να χρησιμοποιεί μια εφαρμογή που λειτουργεί σε ένα εικονικό περιβάλλον της υποδομής μέσα από ειδικές πύλες χωρίς την παρέμβαση άλλων τμημάτων ή υπηρεσιών της υποδομής.

Κάθε εικονικό περιβάλλον είναι ξεχωριστό και περιλαμβάνει ιδιαίτερες τεχνικές απαιτήσεις που καθορίζονται από τον υπεύθυνο για την ανάπτυξη εφαρμογής ενώ κατά την αρχικοποίηση της, καθορίζεται από τον πελάτη το επίπεδο παροχής ποιότητας υπηρεσιών. Αυτές οι απαιτήσεις μεταφέρονται στην υποδομή εικονικού περιβάλλοντος (ISONI) κατά τη διάρκεια της διαπραγμάτευσης των τεχνικών συμφωνιών με τη μεσολάβηση των υπηρεσιών πλαισίου. Να σημειώσουμε ότι η υποδομή εικονικού περιβάλλοντος δεν μπορεί να προσεγγιστεί με άμεσο τρόπο από τους χρήστες (πελάτες ή καταναλωτές) ενώ τα προνόμια πρόσβασής τους περιορίζονται στο επίπεδο υπηρεσιών εφαρμογής.

Οι υπηρεσίες πλαισίου επικοινωνούν συνεχώς και με τους διάφορους τρόπους με την υποδομή εικονικού περιβάλλοντος. Αρχικά κάθε υποδομή εικονικού περιβάλλοντος

δηλώνει τις ικανότητές της στις υπηρεσίες πλαισίου ώστε να μπορεί αναζητηθεί αργότερα και να διαπραγματευτεί τεχνικές συμφωνίες αναφορικά με κάποια εφαρμογή. Επιπλέον η υποδομή εικονικού περιβάλλοντος παρέχει δεδομένα ελέγχου για κάθε εικονικό περιβάλλον στις υπηρεσίες πλαισίου που χρησιμοποιούνται τόσο και το χρόνο εκτέλεσης της εφαρμογής (έλεγχος / control) όσο και κατά το χρόνο σχεδιασμού (ανάπτυξη και μοντελοποίηση / development and modelling). Η λειτουργία σε πραγματικό χρόνο των εφαρμογών και κατά συνέπεια της υποδομής απαιτεί μικρό χρόνο απόκρισης και υψηλές επιδόσεις σε μερικά συστατικά του επιπέδου υπηρεσιών πλαισίου. Για να εξασφαλιστούν λοιπόν αυτά είναι απαραίτητο στιγμιότυπα αυτών των συστατικών να λειτουργούν στο εικονικό περιβάλλον μέσα στο οποίο το επίπεδο παροχής υπηρεσιών (και κατά συνέπεια και η λειτουργία σε πραγματικό χρόνο) είναι εγγυημένο. Όπως είδαμε στο προηγούμενο σχήμα, οι υπηρεσίες πλαισίου αλληλεπιδρούν με τα στιγμιότυπα που εκτελούνται σε ένα εικονικό περιβάλλον με σκοπό τον έλεγχο και την παρακολούθηση της εφαρμογής κατά τη διάρκεια της φάσης εκτέλεσης. Οι υπηρεσίες πλαισίου, τα υποσυστήματα τους και οι διεπαφές περιγράφονται στο επόμενο κεφάλαιο της διατριβής.

Το περιβάλλον εκτέλεσης και τα ευφυή υποσυστήματα δικτύωσης είναι στενά συνδεδεμένα σε αρχιτεκτονικό επίπεδο καθώς βρίσκονται σε συνεχή επικοινωνία κατά τη διάρκεια όλων των διαδικασιών της υποδομής. Αυτά τα υποσυστήματα αποτελούν την υποδομή εικονικού περιβάλλοντος [52]. Ο κύριος στόχος της υποδομής εικονικού περιβάλλοντος είναι η εικονοποίηση των πόρων και η διαχείριση πόρων και υπηρεσιών χωρίς καμία απολύτως γνώση για η ίδια την εφαρμογή.

#### **7.4 Καινοτόμες υπηρεσίες πλαισίου**

Οι υπηρεσίες πλαισίου είναι ένα στρώμα μεταξύ των εφαρμογών και των πόρων που μέσα από την υποδομή εικονικού περιβάλλοντος. Ο βασικός στόχος τους είναι η παροχή και



διαχείριση της εκτέλεση των υπηρεσιών πραγματικού χρόνου υπηρεσιών κατόπιν αιτήσεως του εφαρμογής μέσα σε ένα εικονικό περιβάλλον, προσαρμοσμένο στις ανάγκες της εφαρμογής και των χρηστών όπως αυτά έχουν καθοριστεί στις συμφωνίες εφαρμογής. Εκτός από την εκτέλεση των υπηρεσιών που προσφέρονται στους πελάτες, οι υπηρεσίες πλαισίου υποστηρίζουν τη προσαρμογή νέων εφαρμογών στην υποδομή, τη διαφήμιση υπηρεσιών, την πλήρως αυτοματοποιημένη διαπραγμάτευση SLA, την μετατροπή των υψηλού επιπέδου παραμέτρων εκτέλεσης σε χαμηλό επίπεδο καθώς και την αναζήτηση και δέσμευση πόρων που απαιτούνται για την εκτέλεση μιας εφαρμογής.

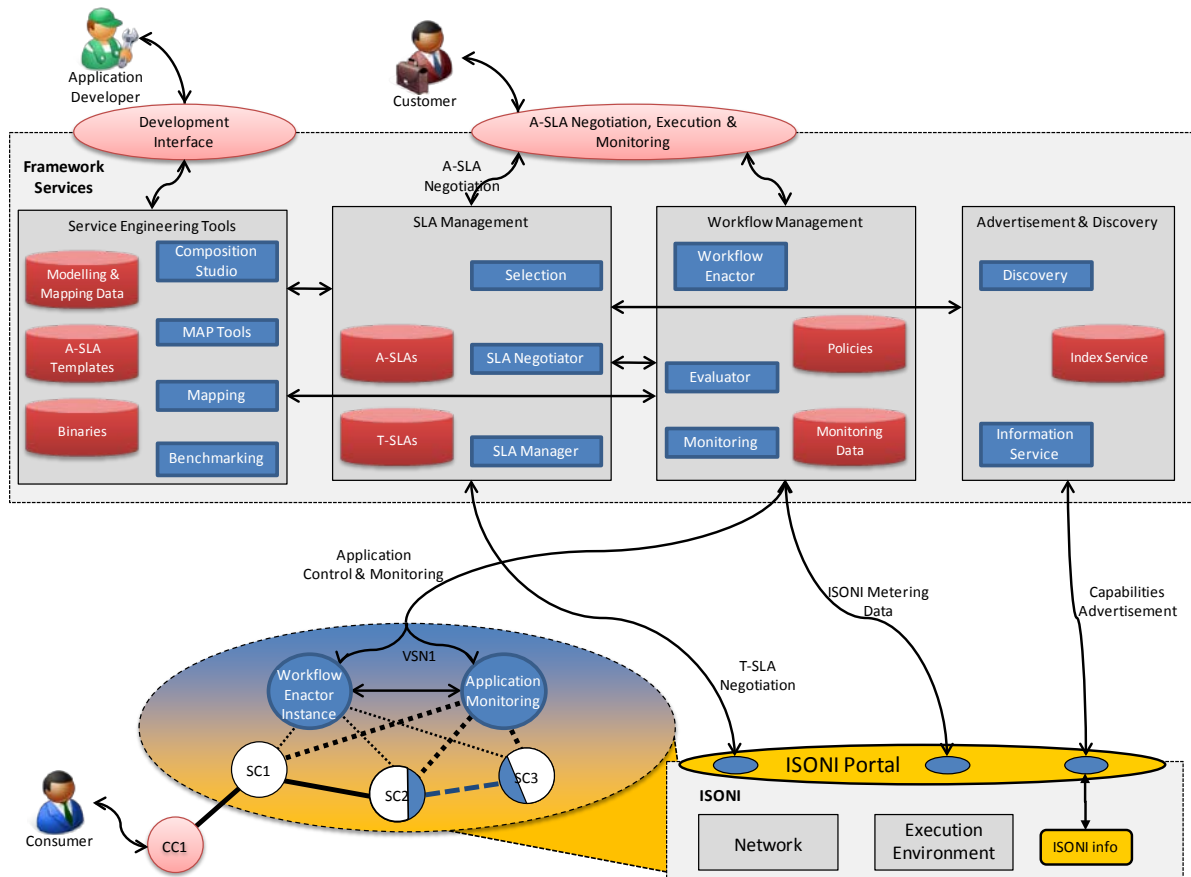
#### **7.4.1 Επισκόπηση υπηρεσιών πλαισίου**

Οι υπηρεσιοστρεφείς μέθοδοι αρχιτεκτονικού σχεδιασμού χρησιμοποιούνται για να επιτευχθεί αυξημένη διαλειτουργικότητα, φορητότητα, ικανότητα επαναχρησιμοποίησης και απλουστευμένος σχεδιασμός. Οι υπηρεσίες πλαισίου εξασφαλίζουν αυτά τα οφέλη ακολουθώντας αυστηρά υπηρεσιοστρεφή αρχιτεκτονικό σχεδιασμό. Το Σχήμα 34 απεικονίζει την αρχιτεκτονική των υπηρεσιών πλαισίου, των κύριων συστατικών και των οντοτήτων και των υποσυστημάτων που με τα οποία αλληλεπιδρούν.

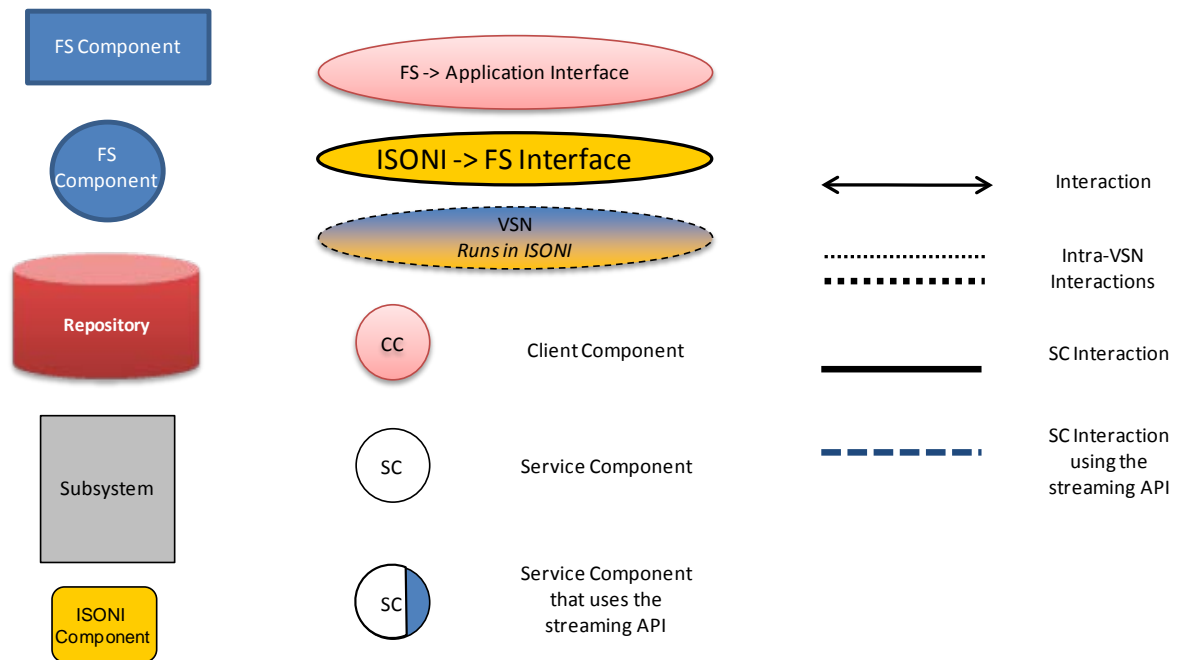
Μερικά στιγμιότυπα υπηρεσιών πλαισίου (π.χ. ο διαχειριστής ροών εργασιών και το σύστημα παρακολούθησης) λειτουργούν μέσα στο εικονικό περιβάλλον για να διαχειριστούν τα στοιχεία της εφαρμογής και για να εγγυηθούν τις σε πραγματικό χρόνο απαιτήσεις ποιότητας υπηρεσιών. Οι υπηρεσίες πλαισίου στηρίζονται στις υπηρεσιοστρεφείς αρχές και αποτελούνται από ένα σύνολο υπηρεσιών, κάθε μία από τις οποίες επιτελεί διαφορετικό έργο, ενώ από την προοπτική του επιπέδου εφαρμογής εμφανίζεται να είναι μία ενιαία υπηρεσία.

Οι υπηρεσίες αυτού του επιπέδου μπορούν να κατηγοριοποιηθούν στα παρακάτω υποσυστήματα ανάλογα με το ρόλο τους. Τα βασικά υποσυστήματα είναι:

- Εργαλεία υποστήριξης εφαρμογής (Service Engineering Tools)
- Διαφήμιση & Αναζήτηση (Advertisement & Discovery)
- Διαχείριση συμφωνιών επιπέδου υπηρεσιών (SLA Management)
- Διαχείριση ροών εργασιών (Workflow Management)



Σχήμα 34: Η αρχιτεκτονική των υπηρεσιών πλαισίου

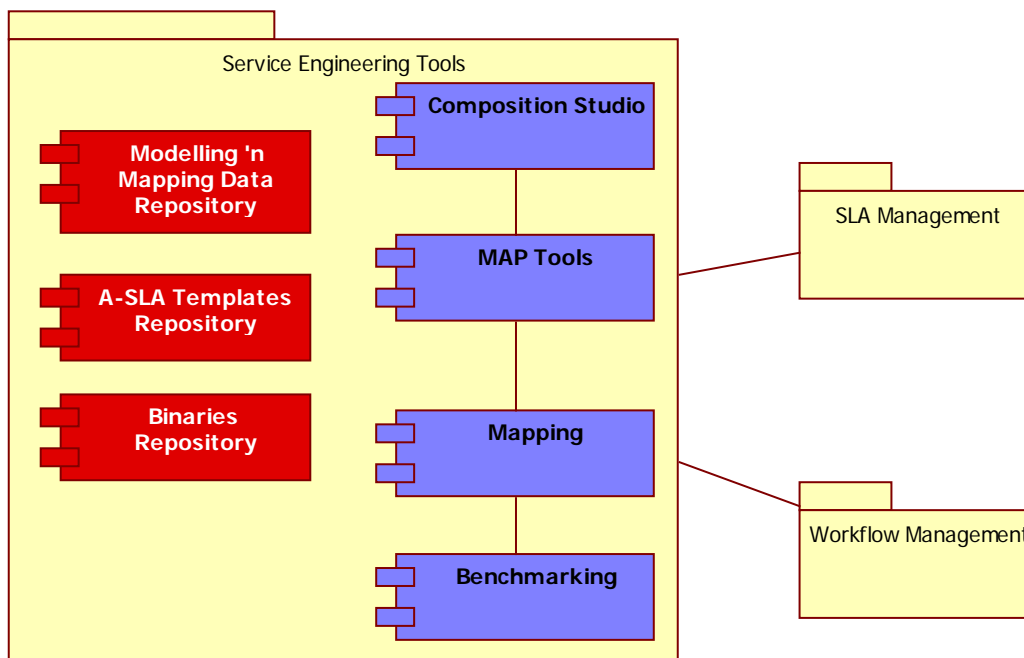


Σχήμα 35: Υπόμνημα αρχιτεκτονικής υπηρεσιών πλαισίου

*Η σελίδα αυτή είναι σκόπιμα λευκή*

#### 7.4.1.1 Εργαλεία υποστήριξης εφαρμογής

Το υποσύστημα εργαλείων υποστήριξης εφαρμογής (Σχήμα 36) είναι αρμόδιο για την προσαρμογή των εφαρμογών στην υποδομή και τη «μετάφραση» των απαιτήσεων της εφαρμογής και των χρηστών σε ένα συγκεκριμένο σύνολο τεχνικών παραμέτρων χαμηλού επιπέδου. Τα εργαλεία υποστήριξης εφαρμογής επικοινωνούν, τόσο κατά τη φάση προσαρμογής της εφαρμογής (offline) όσο και κατά τη φάση εκτέλεσης (online), με τα υποσυστήματα διαχείρισης συμφωνιών επιπέδου υπηρεσιών και διαχείρισης ροών εργασιών.



Σχήμα 36: Εργαλεία υποστήριξης εφαρμογής

Τα δομικά στοιχεία του συγκεκριμένου υποσυστήματος είναι:

- **Στούντιο σύνθεσης (Composition Studio):** Αυτό το συστατικό αποτελεί την κύρια διεπαφή μεταξύ του υπεύθυνου για την ανάπτυξη εφαρμογής και της υποδομής. Μέσω του συστατικού αυτού, ο υπεύθυνος για την ανάπτυξη εφαρμογής μπορεί να δημιουργήσει μια ροή εργασίας από μεμονωμένες υπηρεσίες

*Η σελίδα αυτή είναι σκόπιμα λευκή*

ή ακόμα και να αναπτύξει τις δικές του υπηρεσίες μαζί με όλα τα απαραίτητα μέρη, όπως αρχεία εισόδου, τα πρότυπα SLA, έγγραφα περιγραφής εφαρμογών (Application Service Components Description – ASCD) κλπ. Ο ρόλος των ASCD είναι ιδιαίτερα σημαντικός καθώς περιλαμβάνουν όλες τις παραμέτρους της εφαρμογής καθώς επίσης και τις παραμέτρους για τις οποίες θα «τρέξουν» οι δοκιμαστικές μετρήσεις επιδόσεων.

- **Εργαλεία μοντελοποίησης (MAP Tools):** Τα εργαλεία μοντελοποίησης είναι αρμόδια για τη βελτιστοποίηση και το μετασχηματισμό των αφηρημένων ροών εργασιών που δημιουργούνται από το στούντιο σύνθεσης. Για την διαδικασία αυτή είναι απαραίτητη η συνεργασία με την υπηρεσία αντιστοίχισης για να αποκτήσει πληροφορίες σχετικά με τους χαμηλού επιπέδου πόρους που απαιτούνται για να εξασφαλιστεί το επίπεδο ποιότητας υπηρεσιών της εφαρμογής. Επιπλέον τα εργαλεία μοντελοποίησης χρησιμοποιούνται κατά τη φάση διαπραγμάτευσης συμφωνιών επιπέδου υπηρεσιών και δέσμευσης πόρων για τη δημιουργία των περιγραφών του εικονικού περιβάλλοντος που αποτελούν την βάση της διαπραγμάτευσης των τεχνικών συμφωνιών.
- **Υπηρεσία αντιστοίχισης (Mapping Service):** Αυτή η υπηρεσία είναι αρμόδια για την αντιστοίχιση των υψηλού επιπέδου απαιτήσεων σε παραμέτρους απόδοσης χαμηλού επιπέδου. Συγκεκριμένα, η υπηρεσία αντιστοίχισης δημιουργεί πρότυπα συμπεριφοράς για κάθε μεμονωμένο συστατικό της εφαρμογής με βάση τις μετρήσεις που πραγματοποιήθηκαν κατά την δοκιμαστική λειτουργία (υπηρεσία εκτίμησης επιδόσεων. Οι πληροφορίες αυτές και τα μοντέλα/συναρτήσεις συμπεριφοράς χρησιμοποιούνται από τα εργαλεία μοντελοποίησης προκειμένου να βελτιστοποιηθούν οι ροές εργασιών και να προσδιοριστεί ο ακριβής αριθμός πόρων που απαιτούνται για αυτές.

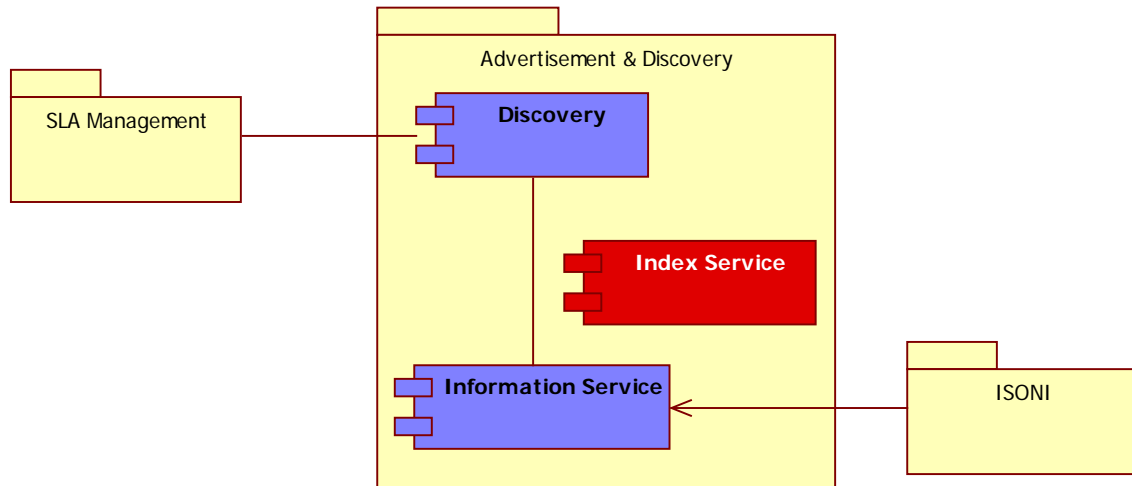
- **Υπηρεσία εκτίμησης επιδόσεων (Benchmarking Service):** Αυτή η υπηρεσία είναι αρμόδια για τη συλλογή των δεδομένων εκτίμησης επιδόσεων των υπηρεσιών της εφαρμογής κατά τη δοκιμαστική λειτουργία της. Τα δεδομένα εκτίμησης επιδόσεων προωθούνται έπειτα στην υπηρεσία αντιστοίχισης όπως περιγράφηκε προηγούμενα.
- **Αποθήκη προτύπων συμφωνιών εφαρμογής (A-SLA Template Repository):** Σε αυτήν την αποθήκη αποθηκεύονται τα πρότυπα συμφωνιών εφαρμογής. Τα πρότυπα αυτά δημιουργούνται από τον πάροχο της εφαρμογής και χρησιμοποιούνται από τον πελάτη βάση των οποίων θα δημιουργήσει την αίτηση SLA κατά τη διαδικασία διαπραγμάτευσης.
- **Αποθήκη λογισμικού υπηρεσιών εφαρμογής (Binaries Repository):** Σε αυτήν την αποθήκη αποθηκεύεται το λογισμικό των υπηρεσιών της εφαρμογής και όλα τα απαραίτητα μέρη (βιβλιοθήκες κ.λπ.). Η υποδομή εικονικού περιβάλλοντος επικοινωνεί με αυτή την αποθήκη πριν την φάση της εκτέλεσης ώστε να αποκτηθεί το απαιτούμενο λογισμικό, να εγκατασταθεί και να αρχικοποιηθεί στο εικονικό περιβάλλον.
- **Αποθήκη μοντέλων και δεδομένων αντιστοίχισης (Modelling & Mapping Data Repository):** Σε αυτήν την αποθήκη διατηρούνται τα μοντέλα και τα δεδομένα αντιστοίχισης για κάθε εφαρμογή και υπηρεσία αντίστοιχα.

#### 7.4.1.2 Διαφήμιση και αναζήτηση

Το υποσύστημα διαφήμισης και αναζήτησης (Σχήμα 37) προσεγγίζεται από τους παρόχους υπηρεσιών και πόρων ώστε αυτοί να διαφημίσουν τις ικανότητές τους ενώ από την άλλη χρησιμοποιείται από το υποσύστημα διαχείρισης συμφωνιών επιπέδου



υπηρεσιών για την αναζήτηση τους με κριτήρια που επιλέγουν οι καταναλωτές της εφαρμογής.



Σχήμα 37: Διαφήμιση και αναζήτηση

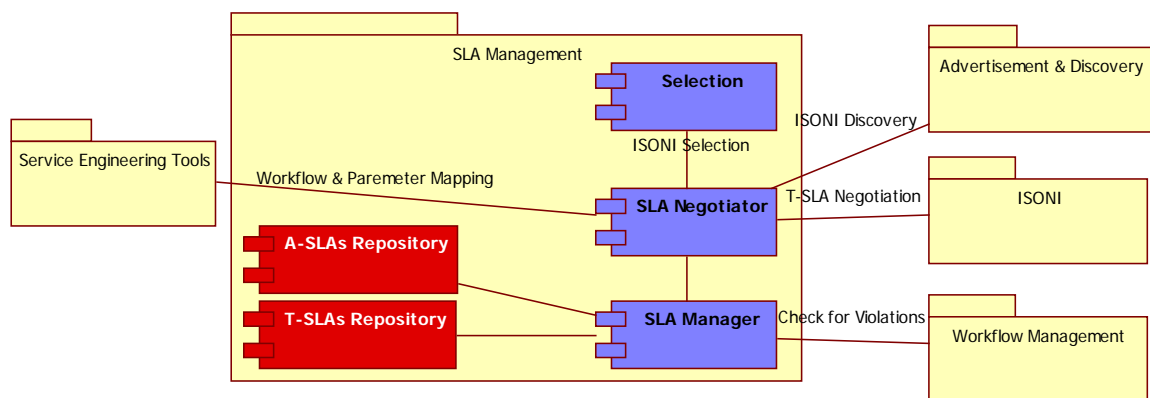
Τα δομικά στοιχεία του συγκεκριμένου υποσυστήματος είναι:

- **Υπηρεσία πληροφοριών (Information and Index Services):** Η υπηρεσία πληροφοριών χρησιμοποιείται από τους παρόχους υπηρεσιών για να διαφημίσουν τις ικανότητές τους προς το επίπεδο των υπηρεσιών πλαισίου. Επομένως, οι υπηρεσίες πλαισίου αγνοούν την ύπαρξη ενός παρόχου εκτός και υπάρχει μια εγγραφή στη συγκεκριμένη υπηρεσία. Οι ικανότητες κάθε παρόχου αποθηκεύονται και χρησιμοποιείται έκτοτε στη διαδικασία αναζήτησης. Αυτές οι εγγραφές περιέχουν τις πληροφορίες για τους τύπους πόρων και της ποιότητας υπηρεσιών που προσφέρει ο κάθε πάροχος.
- **Υπηρεσία αναζήτησης (Discovery Service):** Αυτή η υπηρεσία είναι αρμόδια για την εύρεση των κατάλληλων παρόχων οι οποίοι μπορούν να ικανοποιήσουν τους περιορισμούς που καθορίζονται στην υπό διαπραγμάτευση τεχνική συμφωνία.

*Η σελίδα αυτή είναι σκόπιμα λευκή*

### 7.4.1.3 Διαχείριση συμφωνιών επιπέδου υπηρεσιών

Το υποσύστημα διαχείρισης συμφωνιών επιπέδου υπηρεσιών (Σχήμα 38) υλοποιεί όλες τις διαδικασίες της υποδομής που είναι σχετικές με τα SLAs. Είναι αρμόδιο για τη δυναμική και αυτοματοποιημένη διαπραγμάτευση SLAs τόσο σε επίπεδο συμφωνιών εφαρμογής όσο και σε επίπεδο τεχνικών συμφωνιών. Οι κύριες αλληλεπιδράσεις του υποσυστήματος είναι με τα εργαλεία υποστήριξης της εφαρμογής προκειμένου να αποκτηθούν οι ροές εργασιών και τα πρότυπα διαπραγμάτευσης, με τη σύστημα διαφήμισης & αναζήτησης ώστε να βρεθούν πάροχοι υπηρεσιών με τους οποίους θα διαπραγματευτούν τεχνικές συμφωνίες και τέλος με τη διαχείριση ροών εργασιών για την αξιολόγηση των παραβιάσεων των SLAs.



Σχήμα 38: Διαχείριση συμφωνιών επιπέδου υπηρεσιών

Τα δομικά στοιχεία του συγκεκριμένου υποσυστήματος είναι:

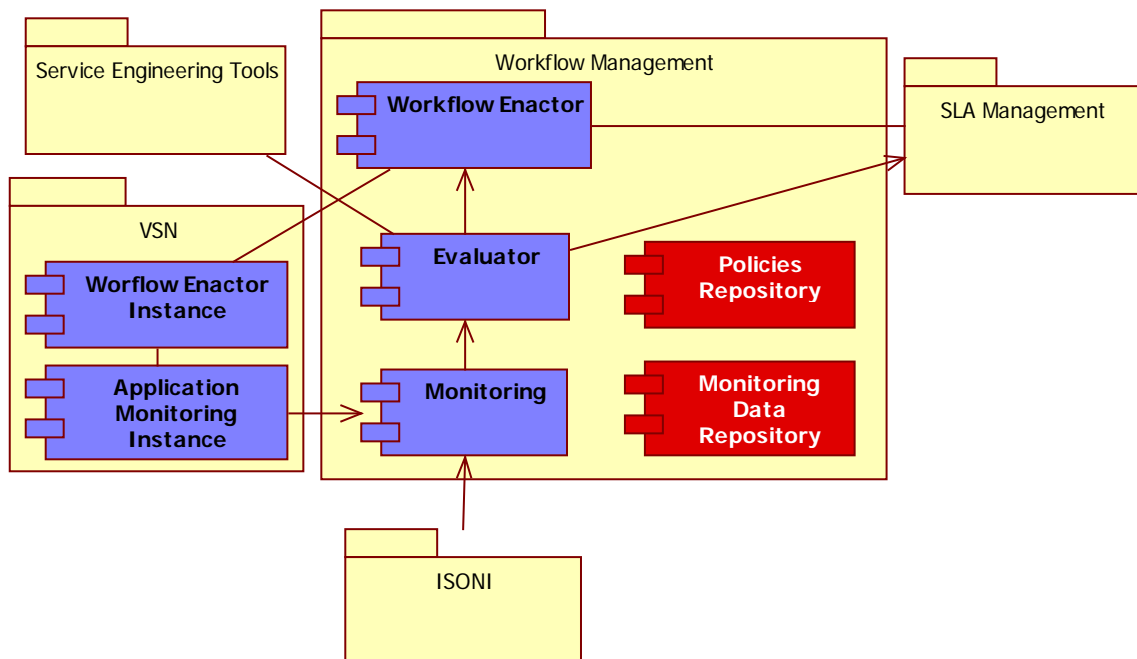
- **Διαπραγματευτής συμφωνιών επιπέδου υπηρεσιών (SLA Negotiator):** Αυτή η υπηρεσία είναι αρμόδια για την επεξεργασία των αιτημάτων παροχής υπηρεσιών και την διαπραγμάτευση SLAs. Είναι επίσης υπεύθυνη για τη δέσμευση των πόρων υπηρεσιών και δικτύου που απαιτούνται για την εκτέλεση των εφαρμογών.

*Η σελίδα αυτή είναι σκόπιμα λευκή*

- **Υπηρεσία επιλογής (Selection Service):** Αυτή η υπηρεσία είναι αρμόδια για την ταξινόμηση και το φιλτράρισμα της λίστας των υποψηφίων παρόχων υπηρεσιών σύμφωνα με τις προτιμήσεις του πελάτη.
- **Διαχειριστής συμφωνιών επιπέδου υπηρεσιών (SLA Manager):** Αυτή η υπηρεσία δημοσιεύει, δημιουργεί και ενημερώνει τα SLA και για αυτόν τον σκοπό, αλληλεπιδρά με τις σχετικές αποθήκες.
- **Αποθήκες συμφωνιών εφαρμογής και τεχνικών συμφωνιών (A-SLA & T-SLA Repositories):** Αυτές οι αποθήκες αποθηκεύουν τις συμφωνίες εφαρμογής καθώς και τις τεχνικές συμφωνίες και έτσι θα πρέπει να είναι προσεγγίσιμα από όλα τα συμβαλλόμενα μέρη με τα οποία το SLA υπογράφεται από κοινού.

#### 7.4.1.4 Διαχείριση ροών εργασιών

Το υποσύστημα διαχείρισης ροών εργασιών (Σχήμα 39) είναι αρμόδιο για την εκτέλεση και τον έλεγχο της εφαρμογής στο επίπεδο υπηρεσιών πλαισίου.



Σχήμα 39: Διαχείριση ροών εργασιών

*Η σελίδα αυτή είναι σκόπιμα λευκή*

Σε αυτό το σημείο πρέπει να σημειωθεί ότι οι καθοριστικές, από άποψη πραγματικού χρόνου απόκρισης, υπηρεσίες πλαισίου όπως η διαχείριση ροών εργασιών και η υπηρεσία παρακολούθησης είναι διαφορετικές για κάθε εικονικό περιβάλλον και αφιερώνονται αποκλειστικά στην υποστήριξη της εκτέλεσης αυτού του συγκεκριμένου εικονικού περιβάλλοντος. Έτσι σε κάθε εικονικό περιβάλλον φιλοξενείται ένα ανεξάρτητο στιγμιότυπο των παραπάνω υπηρεσιών που ελέγχονται όμως από μια υψηλότερου επιπέδου υπηρεσία που βρίσκεται στην περιοχή του παρόχου της υποδομής.

Τα δομικά στοιχεία του συγκεκριμένου υποσυστήματος είναι:

- **Διαχειριστής ροών εργασιών (Workflow Enactor):** Αυτή η υπηρεσία είναι υπεύθυνη για τη σύνθεση των υπηρεσιών που περιλαμβάνονται στις ροές εργασιών της εφαρμογής. Η υπηρεσία διαχείρισης ροών εργασιών ακολουθεί μια ιεραρχική δομή η οποία περιλαμβάνει έναν κεντρικό διαχειριστή και διάφορα στιγμιότυπα του, ένα για κάθε εικονικό περιβάλλον. Ο κεντρικός διαχειριστής βρίσκεται στην περιοχή του παρόχου της υποδομής και είναι αρμόδιος για τον έλεγχο των εκτελέσεων ροών εργασιών σε γενικό επίπεδο ενώ τα στιγμιότυπα που βρίσκονται σε κάθε εικονικό περιβάλλον είναι αρμόδια για την ενορχήστρωση των ροών εργασιών στο επίπεδο του συγκεκριμένου εικονικού περιβάλλοντος. Αυτά τα στιγμιότυπα είναι σε θέση να εκτελέσουν, να σταματήσουν, να παγώσουν, και να επαναλάβουν την εκτέλεσή των υπηρεσιών εφαρμογής που έχουν εγκατασταθεί στο εικονικό περιβάλλον.
- **Υπηρεσία παρακολούθησης (Monitoring):** Αυτή η υπηρεσία ελέγχει την κατανάλωση πόρων κατά τη διάρκεια της εκτέλεσης των υπηρεσιών εφαρμογής. Συγκεντρώνει πληροφορίες τόσο για τις χαμηλού επιπέδου παραμέτρους απόδοσης που προέρχονται από την υποδομή που φιλοξενεί τους πόρους όσο και για τις

υψηλού επιπέδου παραμέτρους απόδοσης που προέρχονται από τα τμήματα υπηρεσιών εφαρμογής που εκτελούνται. Τα δεδομένα παρακολούθησης αποθηκεύονται έτσι ώστε μπορεί να επαναχρησιμοποιηθούν για λόγους βελτιστοποίησης. Πρέπει να σημειωθεί ότι η υπηρεσία παρακολούθησης είναι μία κατανεμημένη υπηρεσία που ακολουθεί ιεραρχική αρχιτεκτονική, ομοίως με το διαχειριστή ροών εργασιών. Συγκεκριμένα, η υπηρεσία παρακολούθησης που βρίσκεται στην περιοχή του παρόχου της υποδομής λειτουργεί ως κεντρικός διαχειριστής ο οποίος συλλέγει δεδομένα από όλα τα στιγμιότυπα υπηρεσιών παρακολούθησης που εκτείνονται σε διαφορετικά εικονικά περιβάλλοντα.

- **Υπηρεσία εκτίμησης (Evaluator):** Όλες πληροφορίες παρακολούθησης θα πρέπει να φιλτραριστούν και θα διαβιβαστούν στην υπηρεσία εκτίμησης που είναι αρμόδια για την εκτέλεση ενός ενεργειών απαιτούνται για τα συμβάντα που καταγράφηκαν ακολουθώντας προκαθορισμένους κανόνες/πολιτικές. Αυτές οι πολιτικές, είναι προκαθορισμένες και μπορούν να είναι είτε συγκεκριμένες για κάθε πολιτική, δημιουργημένες από τον υπεύθυνο για την ανάπτυξη εφαρμογής, είτε να είναι γενικές ακολουθώντας τα επιχειρηματικά μοντέλα των παρόχων της υποδομής.

#### **7.4.2 Φάσεις λειτουργίας**

Η λειτουργία των υπηρεσιών πλαισίου μπορεί να ταξινομηθεί στις ακόλουθες φάσεις:

##### **7.4.2.1 Offline – Προσαρμογή εφαρμογής**

Ο βασικός στόχος αυτής της φάσης είναι προσαρμογή της εφαρμογής στην υποδομή και η δημιουργία των προτύπων SLAs που θα χρησιμοποιηθούν κατά τη διαπραγμάτευση υπηρεσιών. Αρχικά ο πάροχος της εφαρμογής αξιολογεί τη λειτουργία των υπηρεσιών της εφαρμογής στο εικονικό περιβάλλον εκτέλεσης της εφαρμογής. Τα αποτελέσματα αξιολόγησης/επιδόσεων θα χρησιμοποιηθούν στη συνέχεια από την υπηρεσία



αντιστοίχισης για τη δημιουργία προτύπων/συναρτήσεων συμπεριφοράς για κάθε μία υπηρεσία της εφαρμογής. Με αυτά τα πρότυπα συμπεριφοράς ο πάροχος της εφαρμογής είναι σε θέση να δημιουργήσει πρότυπες συμφωνίες επιπέδου υπηρεσιών που θα χρησιμοποιηθούν από τον καταναλωτή της εφαρμογής κατά τη διάρκεια των επόμενων φάσεων.

#### **7.4.2.2 Online – Διαπραγμάτευσης και δέσμευση υπηρεσιών**

Κατά τη διάρκεια αυτής της φάσης, οι υπηρεσίες πλαισίου είναι υπεύθυνες για τη διαπραγμάτευση συμφωνιών επιπέδου υπηρεσιών (εφαρμογής και τεχνικές) και τη δέσμευση των πόρων που απαιτούνται για την εκτέλεση της εφαρμογής. Συγκεκριμένα, ο πελάτης θα επιλέξει και θα συμπληρώσει τις επιθυμητές παραμέτρους ποιότητας υπηρεσιών σε ένα από τα πρότυπα SLA που δημιουργήθηκαν στην προηγούμενη φάση. Αυτές οι υψηλού επιπέδου παράμετροι θα συσχετισθούν στις χαμηλού επιπέδου παραμέτρους χρησιμοποιώντας τα αποτελέσματα της προηγούμενης φάσης για κάθε υπηρεσία της εφαρμογής και συνέχεια θα αναζητηθούν κατάλληλοι πάροχοι πόρων. Για κάθε ένα από τους υποψήφιους παρόχους, θα συνταχθούν και θα διαπραγματευθούν τεχνικές συμφωνίες σύμφωνα με το πρότυπο του παρόχου. Μετά από την επιλογή του πιο κατάλληλου παρόχου ή τεχνική συμφωνία γίνεται επίσημη και οι απαιτούμενοι πόροι δεσμεύονται.

#### **7.4.2.3 Online – Εκτέλεσης και παρακολούθηση**

Σε αυτήν την φάση, οι υπηρεσίες πλαισίου είναι αρμόδιες για την εκτέλεση και τον έλεγχο των εφαρμογών και συγκεκριμένα των ροών εργασιών οι οποίες περιγράφουν επακριβώς τις λειτουργίες τους στη γλώσσα της υποδομής. Ο πελάτης κάθε στιγμή μπορεί να λάβει τα δεδομένα ελέγχου της εκτέλεσης κάθε υπηρεσίας της εφαρμογής στο πλαίσιο της ροής εργασίας. Επίσης οποιαδήποτε στιγμή κατά τη διάρκεια της εκτέλεσης, ο χρήστης εφαρμογής έχει τη δυνατότητα να σταματήσει και να επανεκκινήσει την εκτέλεση των

ροών εργασιών καθώς επίσης και να επαναδιαπραγματευθεί τις συμφωνίες επιπέδου υπηρεσιών εάν το επιθυμεί. Σε περιπτώσεις εξαιρέσεων ή/και παραβιάσεων, οι υπηρεσίες πλαισίου διαβιβάζουν αυτές τις σχετικές πληροφορίες στον καταναλωτή ώστε να ξεκινήσει ή όχι κατά την κρίση του μια διαδικασία επαναδιαπραγμάτευσης.

### **7.4.3 Διεπαφές υπηρεσιών πλαισίου**

Οι υπηρεσίες πλαισίου επικοινωνούν και με ISONI και με το στρώμα εφαρμογής. Οι διεπαφές περιγράφονται λεπτομερώς στις ακόλουθες παραγράφους.

#### **7.4.3.1 Προς την υποδομή εικονικού περιβάλλοντος**

Η επικοινωνία μεταξύ των υπηρεσιών πλαισίου και της υποδομής εικονικού περιβάλλοντος υλοποιείται μέσω μια πύλης που λειτουργεί ως single-point πρόσβασης στην υποδομή. Πρέπει να σημειωθεί ότι εφόσον οι υπηρεσίες πλαισίου στηρίζονται σε τεχνολογίες πλέγματος (Grid) και υπηρεσιών διαδικτύου (Web Services), το SOAP πάνω από HTTP/HTTPS επιλέχθηκε ως πρωτόκολλο επικοινωνίας προκειμένου «να γίνουν κατανοητά» τα μηνύματα που προέρχονται από τις υπηρεσίες πλαισίου προς την υποδομή εικονικού περιβάλλοντος και αντίστροφα. Οι υπηρεσίες πλαισίου χρησιμοποιούν τις διαδικασίες που προβλέπονται από την πύλη προκειμένου να αλληλεπιδράσουν με τις διάφορες υπηρεσίες που βρίσκονται στην υποδομή. Συγκεκριμένα, οι υπηρεσίες πλαισίου επικοινωνούν μέσω της πύλης με τις ακόλουθες υπηρεσίες της υποδομής:

- **Σύστημα πληροφοριών (Information System):** Το σύστημα πληροφοριών υποστηρίζει τη διαδικασία διαφημίσεων των δυνατοτήτων της υποδομής. Ο μηχανισμός διαφημίσεων είναι εύκαμπτος και δυναμικός και οι διαφημίσεις που καταχωρούνται στο επίπεδο υπηρεσιών πλαισίου ενημερώνονται αυτόματα για τις αλλαγές στις δυνατότητες της υποδομής.

- **Διαχείριση συμφωνιών επιπέδου υπηρεσιών (SLA Manager):** Οι υπηρεσίες πλαισίου επικοινωνούν με τον SLA Manager για να διαπραγματευτούν τεχνικές συμφωνίες και να διαχειριστούν τυχόν παραβιάσεις. Η υπηρεσία αυτή ελέγχει εάν υπάρχουν διαθέσιμοι πόροι που να ικανοποιούν το συγκεκριμένο αίτημα στο ζητούμενο χρονικό διάστημα. Εάν υπάρχουν διαθέσιμοι πόροι πραγματοποιείται δέσμευσή τους και αποκρίνεται θετικά στο αίτημα τεχνικής συμφωνίας.
- **Διαχειριστής λειτουργίας (Deployment Manager):** Στο πλαίσιο της παρακολούθησης της εκτέλεσης ροών εργασιών από τις υπηρεσίες πλαισίου απαιτούνται πληροφορίες χαμηλού επιπέδου αναφορικά με την χρήση των κατανεμημένων πόρων. Προκειμένου να εκπληρωθεί αυτός ο στόχος, οι μηχανισμοί παρακολούθησης αλληλεπιδρούν με το διαχειριστή λειτουργίας της υποδομής εικονικού περιβάλλοντος.

#### **7.4.3.2 Προς το επίπεδο των εφαρμογών**

Αν και οι υπηρεσίες πλαισίου είναι μια σύνθεση των διάφορων υπηρεσιών, όπως περιγράφηκε παραπάνω, από την προοπτική του πελάτη εμφανίζονται σαν να είναι μία ενιαία υπηρεσία. Συγκεκριμένα, μια υπηρεσία πυλών υφίσταται μπροστά από τις υπηρεσίες πλαισίου, οι οποίες εξυπηρετούν ως single-point πρόσβασης στους πελάτες εφαρμογής, «κρύβοντας» ταυτόχρονα τις λεπτομέρειες υλοποίησης και την πολυπλοκότητά των υπηρεσιών και της υποδομής. Αυτή η πύλη είναι επομένως αρμόδια για την ερμηνεία των αιτημάτων των πελατών, δρομολόγησή τους στην κατάλληλη υπηρεσία και επιστροφή των αποτελεσμάτων. Η πύλη υπηρεσιών πλαισίου θα υλοποιεί τις ακόλουθες δύο διαδικασίες:

- **Διαπραγμάτευση συμφωνιών εφαρμογής:** Ο πελάτης της εφαρμογής χρησιμοποιήσει την πύλη υπηρεσιών πλαισίου προκειμένου να

(επανα)διαπραγματευτούν συμφωνίες επιπέδου υπηρεσιών και να επαναδιαπραγματευθούν και να δεσμευθούν οι απαιτούμενοι πόροι στην υποδομή εικονικού περιβάλλοντος.

- **Εκτέλεση & Παρακολούθηση της εφαρμογής:** Ο πελάτης εφαρμογής μπορεί να αρχίσει, να σταματήσει και να συνεχίσει την εκτέλεση της ροών εργασιών, καθώς και να ενημερώνεται για τις παραβιάσεις και τα αποτελέσματα παρακολούθησης της εφαρμογής.

## **7.5 Συμπεράσματα**

Σε αυτό το κεφάλαιο παρουσιάσαμε την αρχιτεκτονική μιας νέας υπηρεσιοστρεφούς πλατφόρμας προσανατολισμένη για εφαρμογές πραγματικού χρόνου. Οι δυνατότητες της υποδομής αναφορικά με τις συμφωνίες επιπέδου υπηρεσιών και τη διασφάλιση της ποιότητας υπηρεσιών σε κάθε επίπεδο αλλά και συνολικά αναλύθηκαν εκτενώς. Επιπρόσθετα παρουσιάσαμε τα κύρια υποσυστήματα και τις διεπαφές τους. Τέτοιου είδους υποδομές που συνδυάζουν υπηρεσιοστρεφείς υποδομές και εικονικά περιβάλλοντα αποτελούν από τις πιο κύριες ερευνητικές και επιχειρηματικές τάσεις των ημερών και όλο και περισσότερες λύσεις παρουσιάζονται καθημερινά για να καλύψουν τις αυξανόμενες ανάγκες των χρηστών για δικτύωση και ποιότητα υπηρεσιών.

# 8

## ***Καινοτόμος ενσωμάτωση εφαρμογών βιοϊατρικής σε περιβάλλον πλέγματος***

Για την εκμετάλλευση των τεχνολογιών πλέγματος από την εφαρμογή της *inSilico* ογκολογίας, χρειάζεται προσαρμογή της εν λόγω εφαρμογής στα δεδομένα του περιβάλλοντος του πλέγματος. Στα πλαίσια της διατριβής πραγματοποιήθηκε μια σειρά αλλαγών τόσο στην εφαρμογή, όσο και στο περιβάλλον πλέγματος [95][94][27][25][26]. Καταρχήν, απαιτήθηκαν αλλαγές και τροποποιήσεις του υπάρχοντος πηγαίου κώδικα ώστε να είναι δυνατή η εκτέλεσή του σε περιβάλλον Linux, το οποίο είναι εγκατεστημένο σε όλους τους κόμβους της υποδομής πλέγματος του EGEE [32]. Επιπλέον, ο κώδικας τροποποιήθηκε κατάλληλα, ώστε να είναι εξωτερικά παραμετροποιήσιμος και να είναι δυνατή η αυτόματη εκτέλεση διαφορετικών σεναρίων στο πλέγμα. Η προσαρμογή της *inSilico* Oncology [23][85][89][90] εφαρμογής στο πλέγμα περιέλαβε την ανάπτυξη των κατάλληλων scripts και αρχείων περιγραφής εργασιών, τα οποία είναι απαραίτητα για την εκτέλεση της εφαρμογής στο πλέγμα. Για τις ανάγκες της προσαρμογής της εφαρμογής της υπολογιστικής ογκολογίας σε περιβάλλον πλέγματος [35][34][6][88] κρίθηκε απαραίτητη η δημιουργία μιας δικτυακής πύλης η οποία θα εκθέτει τις υπηρεσίες πλέγματος στους τελικούς χρήστες. Με αυτόν τον τρόπο οι χρήστες της συγκεκριμένης εφαρμογής θα μπορούν να έχουν πρόσβαση σε περιβάλλοντα πλέγματος εύκολα, γρήγορα και αποδοτικά.

Η υποβολή εργασιών στο περιβάλλον πλέγματος του HellasGrid [45] είναι απαραίτητη η χρήση ενός ενδιάμεσου εξυπηρετητή (UI – User Interface). Ο εξυπηρετητής αυτός έχει εγκατεστημένα όλα τα απαραίτητα πακέτα λογισμικού προκειμένου να μπορεί να επικοινωνεί με τους πόρους του πλέγματος (υπολογιστικούς, αποθηκευτικούς και διαχείρισης) ώστε να τους δεσμεύει για την εκτέλεση των εργασιών, να τις υποβάλει και να τις παρακολουθεί (monitoring). Στα αρχικά στάδια δοκιμών μεταφοράς της εφαρμογής σε περιβάλλον πλέγματος χρησιμοποιήθηκε το UI του Δημόκριτου (ui01.isabella.gnet.gr).

Κατά συνέπεια κάθε διεπαφή για αποστολή εργασιών στις υποδομές του HellasGrid πρέπει να επικοινωνεί με κάποιο UI ή να το επεκτείνει και να στηθεί πάνω σε αυτό [14]. Η βέλτιστη λύση για την δημιουργία ενός web portal, που θα καλύπτει τις ανάγκες υποβολής εργασιών σε πλέγμα για την εφαρμογή της Υπολογιστικής Ογκολογίας, είναι να εγκατασταθεί πάνω σε ένα UI εξυπηρετητή και αξιοποιεί τα πακέτα λογισμικού που περιλαμβάνει.

### ***8.1 Ποιότητα υπηρεσίας σε περιβάλλον πλέγματος.***

Ξεκινώντας την ένταξη μια εφαρμογής σε περιβάλλον πλέγματος, εκτός από την ήδη υπάρχουσα πρόοδο στην τεχνολογία Grid, πρέπει να λάβουμε υπόψη τις ανάγκες των χρηστών. Κατά πρώτον, με την χρησιμοποίηση τέτοιων τεχνολογιών, ο χρήστης στοχεύει στην βέλτιστη χρησιμοποίηση των πόρων που ήδη έχει. Αυτό μπορεί να επιτευχθεί με μηχανισμούς προχωρημένης κράτησης, πρόβλεψη του υπολογιστικού φόρτου της εφαρμογής κ.λπ. Ακόμη, μπορεί να αυξήσει το μέγιστο δυναμικό των πόρων του εκμεταλλευόμενος άλλους εξωτερικούς πόρους. Επιπρόσθετα, αυξάνεται η δυνατότητα πρόσβασης των εφαρμογών, από τη μετατροπή των βασικών συστατικών σε Grid-βασισμένες υπηρεσίες που μπορούν να ενσωματωθούν στις υψηλότερου επιπέδου

επιχειρησιακές διαδικασίες ευκολότερα, και βελτιώνεται η όλη διαχειριστική διαδικασία, μέσω της μεγαλύτερης αξιοπιστίας και υπευθυνότητας που παρέχεται[18][49][10].

Η κυριότερη ίσως ανάγκη από την πλευρά του χρήστη οποιουδήποτε υπολογιστικού συστήματος είναι η διεκπεραίωση των απαιτήσεών του σε όσο το δυνατόν πιο λίγο χρόνο, ή σε συγκεκριμένο χρόνο, έτσι ώστε να μπορεί να προγραμματίσει τα καθήκοντά του με τον βέλτιστο δυνατό τρόπο. Δυστυχώς, σε πολλές εφαρμογές αυτό δεν είναι εφικτό, είτε γιατί είναι αρκετά δύσκολη η πρόβλεψη του χρόνου διεκπεραίωσης της εφαρμογής, είτε γιατί αυτό δεν έχει υλοποιηθεί από τον προγραμματιστή της εφαρμογής. Σε τέτοιες περιπτώσεις, ο χρήστης πρέπει να υπολογίζει εμπειρικά τον απαιτούμενο υπολογιστικό χρόνο για την διεκπεραίωση μιας δουλειάς, με αποτέλεσμα συνήθως την μη βελτιστοποιημένη χρήση των πόρων του, λόγω λανθασμένων προβλέψεων. Επιπλέον, ακόμη και στην περίπτωση που ο χρήστης θα μπορούσε να προβλέψει με επάρκεια τον υπολογιστικό χρόνο, δεν θα μπορούσε να τον αλλάξει, για παράδειγμα αν μία συγκεκριμένη δουλειά σε ένα συγκεκριμένο πόρο απαιτεί 1 ώρα υπολογισμών, ο χρήστης δεν θα μπορούσε να απαιτήσει την ολοκλήρωση της δουλειάς σε μισή ώρα, παρά μόνον αν είχε έναν άλλον υπολογιστικό πόρο με τα κατάλληλα στοιχεία. Συνεπώς, η κυριότερη ανάγκη που μπορούν να καλύψουν τα διανεμημένα υπολογιστικά συστήματα είναι η απαίτηση για την ολοκλήρωση μιας δουλειάς σε προδιαγεγραμμένο χρόνο.

Η ποιότητα της υπηρεσίας σε σχέση με τον χρόνο απόκρισης μπορεί να οριστεί με την ικανότητα που έχει το σύστημα να διεκπεραιώσει την αίτηση του χρήστη. Ένα βασικό θέμα για την ποιότητα της υπηρεσίας που ένα τέτοιο σύστημα πρέπει να παρέχει, είναι η επιλογή από το  $n$  χρήστη αντιγράφων της δουλειάς του ώστε εάν η μία εκτέλεση δεν ολοκληρωθεί λόγω κάποιου λάθους, να υπάρχει η επιλογή της ανάκτησης των αποτελεσμάτων από την δεύτερη εκτέλεση.

Η επίτευξη της αίτησης για την ολοκλήρωση μιας δουλειάς σε προδιαγεγραμμένο χρόνο απαιτεί από την αρχιτεκτονική του συστήματος κάποιες βασικές προϋποθέσεις. Χρειάζεται προσεκτική μελέτη της διαδικασίας εκτέλεσης μιας εφαρμογής έτσι ώστε να καθοριστούν οι παράμετροι εκείνοι των πόρων που επηρεάζουν το χρόνο εκτέλεσης. Στη συνέχεια πρέπει να καθοριστούν οι παράμετροι της εφαρμογής που επηρεάζουν το χρόνο εκτέλεσης. Μετά, απαιτείται ο καθορισμός ενός μοντέλου βάσει του οποίου θα υπολογίζεται ο χρόνος εκτέλεσης, με δεδομένες τις παραμέτρους της εφαρμογής και του πόρου στον οποίο πρόκειται να εκτελεστεί. Η ύπαρξη ενός δρομολογητή που θα χρησιμοποιεί το παραπάνω μοντέλο για την εξασφάλιση του ζητούμενου χρόνου απόκρισης είναι απαραίτητη. Η απαιτούμενη αρχιτεκτονική για την διασφάλιση της ζητούμενης ποιότητας υπηρεσίας όπως υλοποιήθηκε στο πρόγραμμα GRID-APP [44] για την εφαρμογή της υπολογιστικής ογκολογίας περιγράφεται στο κεφάλαιο 8.5.

### **8.1.1 Βασικά στοιχεία αρχιτεκτονικής για τη διασφάλιση ποιότητας υπηρεσιών**

Το βασικό στοιχείο που χρειάζεται για την επίτευξη της τήρησης του προδιαγεγραμμένου χρόνου εκτέλεσης είναι το να γνωρίζουμε εκ των προτέρων πόσο χρόνο θα χρειαστεί μια δουλειά για τον υπολογισμό της σε δεδομένη πλατφόρμα. Όσο πιο ακριβής είναι αυτή η εκτίμηση, τόσο καλύτερα μπορούμε να εκμεταλλευτούμε τον χρόνο που διαθέτουν οι πόροι του Grid για υπολογισμό. Για παράδειγμα, εάν ξέρουμε ότι ένας δεδομένος πόρος έχει δέκα λεπτά χωρίς να απασχολείται και επιθυμούμε να τρέξουμε σε αυτόν μια δουλειά που στον συγκεκριμένο πόρο θα διαρκέσει 8 λεπτά, μπορούμε να δρομολογήσουμε την δουλειά εκεί. Αν όμως η εκτίμησή μας είναι λανθασμένη και τελικά η δουλειά για να ολοκληρωθεί χρειάζεται χρόνο μεγαλύτερο των δέκα λεπτών, τότε ο υπολογιστής μην έχοντας ελεύθερους πόρους μπορεί μέχρι και να διακόψει την εκτέλεση οδηγώντας σε αποτυχία. Συνεπώς, είναι αναγκαία η μελέτη των παραμέτρων απόδοσης των πόρων και



των παραμέτρων φόρτου της εφαρμογής, ώστε να κατασκευαστεί ένα γενικό μοντέλο που να μπορεί να καλύψει με αρκετή ακρίβεια τον υπολογισμό του χρόνου εκτέλεσης μιας δεδομένης δουλειάς σε ένα δεδομένο πόρο.

Οι οντότητες που υλοποιούν αυτή την απαίτηση είναι ο Workload Estimator, ο Capacity Estimator και ο Scheduler της εφαρμογής.

- **Capacity Estimator:** Ο εκτιμητής πόρων πρέπει να υπολογίζει τις βασικές παραμέτρους απόδοσης ενός πόρου, ώστε η έξοδός του να αποτελεί μέτρο σύγκρισης της απόδοσης των πόρων. Οι παράμετροι αυτοί πρέπει να είναι ίδιες για κάθε πόρο, έτσι ώστε το σύστημα να χρησιμοποιεί ένα ενιαίο μοντέλο απόδοσης για ετερογενείς πόρους. Με την χρησιμοποίηση ενός τέτοιου μοντέλου οι ετερογενείς πόροι αντιμετωπίζονται από το σύστημα με τον ίδιο τρόπο γιατί ο εκτιμητής πόρων λειτουργεί ως αντικειμενικό μέτρο σύγκρισης της απόδοσής τους.
- **Workload Estimator:** Ο εκτιμητής φόρτου εφαρμογής υπολογίζει το φορτίο που εφαρμόζει η εφαρμογή σε κάποιο πόρο. Συνεπώς, οι παράμετροι που υπολογίζουν τον φόρτο θα πρέπει να συνδέονται άμεσα με τις παραμέτρους του εκτιμητή πόρων, ώστε να δείχνουν επιπλέον ποια στοιχεία της απόδοσης επηρεάζουν το χρόνο εκτέλεσης. Οι παράμετροι φόρτου πρέπει να είναι ίδιες για κάθε εφαρμογή γιατί κάθε διαφορετική εφαρμογή χρησιμοποιεί τους ίδιους υπολογιστικούς πόρους με τις ίδιες παραμέτρους απόδοσης [17].
- **Scheduler:** Ο δρομολογητής (scheduler) αναλαμβάνει την δρομολόγηση των δουλειών στους πόρους βάσει ορισμένων κριτηρίων, όπως για παράδειγμα την επίτευξη του προκαθορισμένου χρόνου εκτέλεσης, χρησιμοποιώντας την εκτίμηση για το χρόνο εκτέλεσης από το παραπάνω μοντέλο [11][16][102][87][104].

## 8.2 Εκτίμηση υπολογιστικού φόρτου

Η εκτίμηση υπολογιστικού φόρτου (Workload Estimation) είναι μια πολύ σημαντική διαδικασία κατά την υποβολή εργασιών σε περιβάλλον πλέγματος. Η εφαρμογή της υπολογιστικής ογκολογίας έχει χαρακτηριστικά τέτοια που διευκολύνουν την διαδικασία εκτίμησης του υπολογιστικού φόρτου.

Κάθε εργασία υπολογιστικής ογκολογίας αποτελείται από μικρότερες εργασίες που εκτελούνται σειριακά. Σε κάθε μονάδα εργασίας τα δεδομένα εισόδου είναι τα ίδια, ενώ το μόνο που διαφοροποιείται σε κάθε εκτέλεση είναι οι παράμετροι εκτέλεσης. Αρχικά σε κάθε υποβολή εργασίας ο χρήστης επιλέγει τον αλγόριθμο εκτέλεσης και τις παραμέτρους της. Να σημειωθεί ότι κάθε παράμετρος μπορεί να έχει περισσότερες από μία τιμές για την ολοκλήρωση της εξομοίωσης. Άρα λοιπόν ο συνολικός αριθμός των μονάδων εργασίας που θα εκτελεστούν είναι οι συνδυασμοί όλων των παραμέτρων εκτέλεσης, δηλαδή:

$$N_{εκτελ.} = \prod_{k=1}^n P_k$$

όπου  $P_k$  το σύνολο των διαφορετικών τιμών για την παράμετρο  $k$  και  $n$  ο συνολικός αριθμός των παραμέτρων. Για παράδειγμα, για τον αλγόριθμο της ακτινοθεραπείας ο οποίος περιλαμβάνει 5 παραμέτρους και αν υποθέσουμε ότι για την συγκεκριμένη εργασία έχουμε 2 διαφορετικές τιμές για 4 από τις 5 παραμέτρους, προκύπτει ότι ο συνολικός αριθμός των μονάδων εργασίας που θα εκτελεστούν είναι  $1 \times 2 \times 2 \times 2 \times 2 = 16$ .

Άλλο πλεονέκτημα για την εφαρμογή της υπολογιστικής ογκολογίας είναι ότι ο χρόνος εκτέλεσης για κάθε μονάδα εργασίας είναι ο ίδιος για κάθε αλγόριθμό. Έτσι λοιπόν ο συνολικός αναμενόμενος χρόνος εκτέλεσης είναι:

$$t_{ολ} = N_{εκτελ.} * t_{αλγ}$$

όπου  $t_{αλγ}$  ο χρόνος εκτέλεσης της μονάδας εργασίας για τον συγκεκριμένο αλγόριθμο.

Ο υπολογισμός του υπολογιστικού φόρτου πραγματοποιείται αυτόματα κατά την έναρξη διαδικασίας υποβολής εργασίας μέσα στην εφαρμογή. Το αποτέλεσμα της εκτίμησης εισάγεται αυτόματα στην διαδικασία δρομολόγησης με στόχο την βελτιστοποίηση της ποιότητας υπηρεσιών. Οι λεπτομέρειες παρουσιάζονται αναλυτικά σε επόμενο κεφάλαιο.

### 8.3 Εκτίμηση απόδοσης πόρων

Η εκτίμηση απόδοσης των υπολογιστικών πόρων (Capacity Estimation) είναι μια διαδικασία απαραίτητη για την αποδοτική δρομολόγηση εργασιών σε περιβάλλον πλέγματος.

Η υποδομή του HellasGrid περιλαμβάνει δυνατότητες εκτίμησης απόδοσης πόρων. Συγκεκριμένα με την εντολή `lcg-infosites` του `gLite` [42] μπορούμε να έχουμε με μεγάλη ακρίβεια πληροφορίες για όλους τους Υπολογιστικούς Πόρους (Computing Elements) της υποδομής.

#CPU	Free	Total Jobs	Running	waiting	ComputingElement
125	1	699	60	639	ce.phy.bg.ac.yu:2119/jobmanager-pbs-see
41	1	750	38	712	ce64.phy.bg.ac.yu:2119/jobmanager-pbs-see
13	3	1	0	1	ce002.ipp.acad.bg:2119/jobmanager-lcgpbs-see
13	3	1	0	1	ce001.ipp.acad.bg:2119/blah-pbs-see
15	7	7	6	1	ce001.ibm.bas.bg:2119/jobmanager-lcgpbs-see
74	56	1	0	1	ce02.grid.acad.bg:2119/jobmanager-pbs-see
74	56	1	0	1	ce01.grid.acad.bg:2119/blah-pbs-see
80	65	4	4	0	ce101.grid.ucy.ac.cy:2119/jobmanager-lcgpbs-see
10	0	8	8	0	ce301.intercol.edu:2119/jobmanager-lcgpbs-see
9	1	3	2	1	node001.grid.auth.gr:2119/jobmanager-pbs-see
32	7	0	0	4444	ce.hep.ntua.gr:2119/jobmanager-lcgpbs-see
8	1	4	0	4	grid001.ics.forth.gr:2119/jobmanager-lcgpbs-see
20	9	5	4	1	ce02.marie.hellasgrid.gr:2119/jobmanager-pbs-see
44	1	8	6	2	ce01.isabella.grnet.gr:2119/jobmanager-pbs-see
118	14	5	5	0	ce01.marie.hellasgrid.gr:2119/jobmanager-pbs-see
118	14	5	5	0	glite-ce01.marie.hellasgrid.gr:2119/blah-pbs-see
116	6	205	45	160	ce01.afrodit1.hellasgrid.gr:2119/jobmanager-pbs-see
116	59	29	29	0	ce01.kallisto.hellasgrid.gr:2119/jobmanager-pbs-see
110	74	4	4	0	ce01.ariadni.hellasgrid.gr:2119/jobmanager-lcgpbs-see
214	128	10	9	1	ce01.athena.hellasgrid.gr:2119/jobmanager-pbs-see
214	128	1	1	0	ce01.athena.hellasgrid.gr:2119/jobmanager-pbs-see
214	128	10	9	1	ce02.athena.hellasgrid.gr:2119/blah-pbs-see
214	128	1	1	0	ce02.athena.hellasgrid.gr:2119/blah-pbs-long
44	28	8	8	0	cs-grid0.bgu.ac.il:2119/jobmanager-lcgpbs-see
44	29	7	7	0	cs-grid1.bgu.ac.il:2119/blah-pbs-see
38	14	20	20	0	grim-ce.iucc.ac.il:2119/jobmanager-lcgpbs-see
29	26	0	0	0	grid-ce.fi.edu.mk:2119/jobmanager-pbs-see
8	3	4	2	2	testbed001.grid.ici.ro:2119/jobmanager-pbs-see
277	191	13	10	3	tbit01.nipne.ro:2119/jobmanager-lcgpbs-see
54	16	3	3	0	lcfgng.cs.tau.ac.il:2119/jobmanager-pbs-see
113	84	12	8	4	ce.ulakbim.gov.tr:2119/jobmanager-lcgpbs-see
308	28	2	2	0	cox01.grid.metu.edu.tr:2119/jobmanager-lcgpbs-see
60	33	0	0	0	grid01.erciyes.edu.tr:2119/jobmanager-lcgpbs-see
64	29	14	14	0	yildirim.grid.boun.edu.tr:2119/jobmanager-lcgpbs-see
60	32	13	13	0	paugrid1.pamukkale.edu.tr:2119/jobmanager-lcgpbs-see
64	14	1	1	0	grid01.cu.edu.tr:2119/jobmanager-lcgpbs-see
64	63	3	1	2	ituce.grid.itu.edu.tr:2119/jobmanager-lcgpbs-see
288	4	2	2	0	kalkan1.ulakbim.gov.tr:2119/jobmanager-lcgpbs-see
87	35	0	0	0	wipp-ce.weizmann.ac.il:2119/jobmanager-lcgpbs-see

Σχήμα 40: Αποτελέσματα εκτίμησης απόδοσης πόρων.

*Η σελίδα αυτή είναι σκόπιμα λευκή*

Οι πληροφορίες που είναι διαθέσιμες με αυτόν τον τρόπο στην εφαρμογή είναι:

- Ο συνολικός αριθμός των επεξεργαστών (CPUs) που διαθέτει το συγκεκριμένο CE,
- αριθμός των διαθέσιμων-ελεύθερων επεξεργαστών,
- Οι συνολικές εργασίες που έχουν υποβληθεί στο CE και πόσες από αυτές τρέχουν ή είναι στην ουρά.

Στην παραπάνω εικόνα διακρίνονται οι πληροφορίες απόδοσης του HellasGrid για μια δεδομένη χρονική στιγμή. Οι πληροφορίες αυτές όπως θα δούμε παρακάτω αξιοποιούνται από το δρομολογητή για να πετύχει αποδοτικότερη εκτέλεση των εργασιών.

### **8.3.1 Καταγραφή επιπέδου ποιότητας υπηρεσιών και αξιοποίηση της**

Κατά την εκτέλεση εργασιών υπολογιστικής ογκολογίας στις υποδομές του HellasGrid διαπιστώσαμε ότι οι πληροφορίες που παρείχε το πλέγμα σχετικά με την απόδοση των πόρων του δεν ήταν ακριβείς. Επιπρόσθετα εργασίες που αποστέλλονταν σε συγκεκριμένους κόμβους αποτύγχαναν πάντα, παρόλο που το πλέγμα θεωρητικά τις υποστήριζε.

Η λύση που δόθηκε στο συγκεκριμένο πρόβλημα ήταν η αποθήκευση πληροφοριών σχετικά με τις υποβληθείσες εργασίες και τους εξυπηρετητές που εκτελέστηκαν στην βάση δεδομένων της εφαρμογής. Με αυτό τον τρόπο πληροφορίες που αφορούν τις εργασίες υπολογιστικής βιολογίας και την ικανότητα των εξυπηρετητών να τις εκτελέσουν αξιοποιούνται για την αποδοτικότερη υποβολή των νέων εργασιών.

Ειδικότερα μετά το τέλος της κάθε εργασίας, επιτυχημένης ή όχι, στην βάση δεδομένων αποθηκεύεται ο χρόνος εκτέλεσης της μονάδας εργασίας και φυσικά η επιτυχημένη ή όχι εκτέλεση. Έτσι λοιπόν ο δρομολογητής χρησιμοποιεί την πληροφορία αυτή για να επιλέξει καλύτερο εξυπηρετητή.

## **8.4 Δρομολόγηση εργασιών**

Η διαχείριση πόρων σε ένα σύστημα Grid έχει την ικανότητα να εγγυηθεί στις εφαρμογές μια εξατομικευμένη ποιότητα υπηρεσίας (QoS). Αυτή η ποιότητα υπηρεσίας (QoS) πρέπει να παρασχεθεί σε ένα συχνά μη-πρόσφορο περιβάλλον, όπου οι πόροι είναι λιγοστοί, και ποικίλλουν κατά διαστήματα κατά έναν δυναμικό τρόπο. Θεωρούμε την ποιότητα υπηρεσίας (QoS) σχεδιάζοντας το πρόβλημα του οποίου ο στόχος είναι να καθορίσει πότε και σε ποιο πόρο μια δεδομένη στοιχειώδης εργασία πρέπει να εκτελεσθεί, προκειμένου να ικανοποιηθούν οι ζητούμενες παράμετροι QoS. Παρακάτω προτείνεται ένας νέος αλγόριθμος για την δρομολόγηση με QoS σε ένα Grid. Ο αλγόριθμος χρησιμοποιεί μια προσέγγιση δίκαιης διανομής για την παροχή δίκαιης πρόσβασης στους πόρους Grid για όλους τους χρήστες.

Η δυνατότητα να παρασχεθεί η προσυμφωνηθείσα ποιότητα της υπηρεσίας (QoS) είναι μεγάλης σημασίας για την επιτυχία του Grid, επειδή εάν δεν μπορεί να παρασχεθεί, οι χρήστες μπορεί να είναι απρόθυμοι να χρησιμοποιήσουν πάλι τις υπηρεσίες Grid.

Ένα σύστημα Grid είναι ουσιαστικά η δυναμική συνάθροιση πόρων από τους προμηθευτές υπηρεσίας, με σκοπό τη διανομή αυτών των πόρων κατά απαίτηση και αποτελεσματικά. Προκειμένου να παρασχεθεί το συμφωνηθέν QoS σε έναν χρήστη, ο διαχειριστής πόρων πρέπει να είναι σε θέση να κρατήσει διαθέσιμα (μέρη από) τους πόρους. Η κατά απαίτηση ζήτηση και η αποδοτική χρήση των πόρων σε ένα τέτοιο περιβάλλον υπονοεί ότι οι πόροι (ή μέρη των πόρων) πρέπει να διατηρηθούν για μια στοιχειώδη εργασία μόνο για τη χρονική περίοδο κατά τη διάρκεια της οποίας χρησιμοποιούνται πραγματικά, και πρέπει να είναι διαθέσιμα σε άλλες στοιχειώδεις εργασίες για τον υπόλοιπο χρόνο. Αυτό οδηγεί στην ανάγκη της χρονομετρημένης κράτησης των πόρων και της δρομολόγησης των στοιχειωδών εργασιών.

Μια επιτυχημένη εφαρμογή Grid πρέπει να έχει χαρακτηριστικά ώστε πολλοί χρήστες να θελήσουν να την χρησιμοποιήσουν. Όμως το πλέγμα μπορεί να έχει περισσότερους χρήστες από όσους είναι σε θέση να χειριστούν οι υπάρχοντες πόροι. Η συνάντηση των απαιτήσεων QoS του ενός χρήστη, εντούτοις, δεν πρέπει να επιτευχθεί με τη θυσία του QoS ενός άλλου χρήστη. Όταν το επιθυμητό QoS δεν μπορεί να παρασχεθεί, η υποβάθμιση στο παρεχόμενο QoS πρέπει να είναι μοιρασμένη δίκαια σε όλους τους χρήστες. Αυτό οδηγεί φυσικά στην ανάγκη του ελέγχου συμφόρησης (congestion control) και τη συνδεμένη έννοια της δικαιοσύνης (fairness), τα οποία είναι δύο από τα ζητήματα που προσπαθούμε να εξετάσουμε εδώ. Η δικαιοσύνη δεν σημαίνει απαραίτητα ότι όλοι οι χρήστες έχουν ίση πρόσβαση στους πόρους Grid.

Όταν δεν υπάρχει καμία συμφόρηση, κάθε χρήστης παίρνει το QoS που ζήτησε, αλλά όταν προκύπτει η συμφόρηση (σημαίνοντας ότι το επιθυμητό QoS δεν μπορεί να παραδοθεί σε όλους τους χρήστες), το QoS που κάθε χρήστης λαμβάνει (όπως μετριέται από την υπολογιστική ισχύ που του δίνεται, ή το χρόνο ολοκλήρωσής του, ή το ποσό χρόνου κατά το οποίο χάνει την προθεσμία του) υποβιβάζεται.

Η δρομολόγηση και η διαχείριση των πόρων, που στοχεύουν στη βελτιστοποίηση της κατανομής των πόρων Grid, είναι σημαντικά ζητήματα στην υλοποίηση μιας υπολογιστικής υποδομής Grid και στον καθορισμό της δυνατότητάς της να παραδώσει το ζητούμενο QoS και να παρέχει τη δίκαιη πρόσβαση σε όλους τους χρήστες. Ο διαχειριστής πόρων λαμβάνει τις πληροφορίες για τα χαρακτηριστικά και τις απαιτήσεις των εργασιών που υποβάλλονται από τους διαφορετικούς χρήστες και καθορίζει πότε και σε ποιο επεξεργαστή θα εκτελεστεί κάθε εργασία. Ο στόχος που θέτουμε για το διαχειριστή πόρων ενός Grid είναι να αναθέτει τους υπολογιστικούς πόρους στις στοιχειώδεις εργασίες με έναν αποδοτικό και δίκαιο τρόπο, συναντώντας στο μέγιστο πιθανό βαθμό τις απαιτήσεις QoS των μεμονωμένων στοιχειωδών εργασιών. Η

αποδοτικότητα στη χρήση των πόρων είναι σημαντική επειδή αυτό είναι το ζητούμενο που παρακίνησε αρχικά την ιδέα για την δημιουργία του Grid. Η δικαιοσύνη είναι σημαντική επειδή είναι έμφυτη στην έννοια της διανομής, και η διανομή των πόρων είναι το κυρίως ζήτημα του Grid. Η συνάντηση των απαιτήσεων των χρηστών είναι σημαντική επειδή εάν το QoS δεν είναι εγγυημένο οι χρήστες δεν θα θελήσουν να χρησιμοποιήσουν. Για να συναντήσει τους στόχους του ο διαχειριστής πόρων πρέπει να λάβει υπόψη τον αναμενόμενο φόρτο εργασίας των εφαρμογών που πρόκειται να δρομολογηθούν, των απαιτήσεων QoS τους (συμπεριλαμβανομένων των προθεσμιών), των καθυστερήσεων επικοινωνίας που περιλαμβάνονται, του τρέχοντος και του μελλοντικού σχεδίου χρησιμοποίησης των διαθέσιμων επεξεργαστών (utilization profile).

#### **8.4.1 Διαχείριση πόρων και δρομολόγηση**

Το κεντρικό στοιχείο οποιασδήποτε αρχιτεκτονικής Grid είναι το διαχειριστικό εργαλείο των πόρων που χρησιμοποιείται για να δρομολογήσει τις στοιχειώδεις εργασίες υπολογισμού και επικοινωνίας στους διαθέσιμους πόρους. Τα εργαλεία μέτρησης απόδοσης, ο χάρτης των πόρων, και οι απαιτήσεις και τα χαρακτηριστικά των στοιχειωδών εργασιών (αναμενόμενη "διάρκεια" στοιχειώδους εργασίας, προθεσμίες στοιχειωδών εργασιών, απαιτήσεις ανοχής λαθών κ.λπ.) διαμορφώνουν τις εισόδους, και οι αποφάσεις σχετικά με την επεξεργασία, την αποθήκευση, και τους πόρους επικοινωνίας που θα χρησιμοποιηθούν, και οι χρόνοι για τους οποίους θα διατηρηθούν διαμορφώνουν τα αποτελέσματα εξόδου του δρομολογητή. Το Grid είναι εγγενώς δυναμικό: φορτία μηχανών, λανθάνουσες καθυστερήσεις δικτύων και εύρη ζώνης δικτύων, αλλάζουν συνεχώς. Οι διαθέσιμοι πόροι και οι απαιτήσεις εφαρμογής μπορούν να είναι εντυπωσιακά διαφορετικά από τη μια ώρα στην άλλη. Οι πόροι δεν αφιερώνονται αποκλειστικά στη χρήση Grid, και μπορούν προσωρινά να αφαιρεθούν από το Grid για να εργαστούν απλώς σε ένα τοπικό πρόβλημα. Οι διαχειριστικοί αλγόριθμοι πόρων πρέπει να



είναι σε θέση να προσαρμοστούν σε τέτοια δυναμικά περιβάλλοντα, κάτι το οποίο απαιτεί νέα σκέψη για τον τρόπο σχεδίασης των αλγορίθμων δρομολόγησης.

Είναι καλά γνωστό από το συμβατικό παράλληλο και διανεμημένο υπολογισμό ότι οι υπολογιστές  $N$  που λειτουργούν σε ένα δεδομένο σύνολο στοιχειωδών εργασιών υπολογισμού δεν τις εκτελούν απαραίτητως  $N$  χρόνο γρηγορότερα από έναν ενιαίο υπολογιστή που λειτουργεί μόνος του.

Ένας (αλλά όχι ο μόνος) λόγος για αυτό είναι η καθυστέρηση επικοινωνίας. Οι υπολογιστές που τρέχουν μια εφαρμογή σε ένα Grid, πρέπει να πάρουν τον απαιτούμενο κώδικα και τα δεδομένα από μια απομακρυσμένη μηχανή. Επίσης, κατά τη διάρκεια των υπολογισμών, οι μηχανές πρέπει συχνά να ανταλλάζουν ενδιάμεσα αποτελέσματα. Σε πολλές περιπτώσεις, ο χρόνος που απαιτείται για την επικοινωνία είναι περισσότερος από το χρόνο που ξοδεύεται από τις μηχανές για τους υπολογισμούς. Εάν τα Grid πρόκειται τελικά να χρησιμοποιηθούν για τους λεπτούς υπολογισμούς (οι ιστορίες επιτυχίας του Grid περιορίζονται κυρίως σε χονδροειδή υπολογισμό), η επικοινωνία θα γίνει μια τροχοπέδη απόδοσης. Είναι επομένως σημαντικό ο δρομολογητής ενός διαχειριστή πόρων Grid να παίρνει υπόψη τις καθυστερήσεις επικοινωνίας κατά την ανάθεση των στοιχειωδών εργασιών στους επεξεργαστές.

#### **8.4.2 Δρομολόγηση στο μεσολογισμικό**

Η δρομολόγηση που ενσωματώθηκε στην εφαρμογή μας είναι σε 2 επίπεδα. Αρχικά είναι στο επίπεδο της δικτυακής πύλης και έπειτα στο πλέγμα με τους μηχανισμούς που ήδη διαθέτει. Παράλληλα όμως, η εφαρμογή δίνει στους χρήστες την δυνατότητα παράκαμψης των μηχανισμών δρομολόγησης και την επιλογή συγκεκριμένων εξυπηρετητών με μη αυτόματο τρόπο.

Κατά την αυτοματοποιημένη διαδικασία ο χρήστης επιλέγει μόνο τον αλγόριθμο εκτέλεσης, εισάγει τα αρχεία που απαιτούνται και καθορίζει τις επιθυμητές παραμέτρους εκτέλεσης. Στη συνέχεια η εφαρμογή αξιολογεί τα στοιχεία που εισήγαγε ο χρήστης και εκτιμά τον φόρτο της συγκεκριμένης εργασίας. Ταυτόχρονα, η εφαρμογή επικοινωνεί με το πλέγμα για να εκτιμήσει την διαθεσιμότητα των υπολογιστικών πόρων στο πλέγμα ενώ εξετάζεται και η ποιότητα υπηρεσιών του κάθε πόρου με βάση τα δεδομένα που υπάρχουν καταχωρημένα από προηγούμενες εκτελέσεις παρόμοιων εργασιών.

Με βάση όλες τις προηγούμενες πληροφορίες πραγματοποιείται η υποβολή της εργασίας στο πλέγμα, είτε κάποιο συγκεκριμένο πόρο, είτε σε μια λίστα από διαθέσιμους που πληρούν όλες τις απαιτήσεις, αφήνοντας την τελική επιλογή στους μηχανισμούς του πλέγματος.

Αφού γίνει η υποβολή κάποιας εργασίας στο πλέγμα, ακολουθεί η διαδικασία του match-making του μεσολογισμικού gLite, η οποία έχει ως στόχο την εύρεση του καταλληλότερου computing element (CE) στο πλέγμα, σύμφωνα με συγκεκριμένα κριτήρια που μπορούν να οριστούν είτε απευθείας από το χρήστη είτε μέσα από το πρώτο επίπεδο δρομολόγηση στην δικτυακή πύλη. Τα κριτήρια αυτά ορίζονται κατά την περιγραφή της εργασίας στο JDL αρχείο, με συγκεκριμένα attributes. Συγκεκριμένα, για τον έλεγχο του match-making των εργασιών μπορούν να χρησιμοποιηθούν τα attributes *Requirements* και *Rank*, τα οποία περιγράφονται στη συνέχεια.

#### **8.4.2.1 Requirements:**

Το attribute αυτό μπορεί να χρησιμοποιηθεί για να εκφραστούν περιορισμοί στους πόρους όπου θα εκτελεστεί η εργασία. Η τιμή του είναι μια boolean έκφραση που πρέπει να είναι αληθής για ένα συγκεκριμένο CE ώστε να εκτελεστεί η εργασία σε αυτό. Για τη σύνθεση αυτών των boolean εκφράσεων μπορούν να χρησιμοποιηθούν όλα τα attributes του GLUE schema [91]. Σε κάθε JDL αρχείο μπορεί να οριστεί μόνο ένα Requirements attribute. Αν

απαιτείται να εφαρμοστούν πολλαπλές συνθήκες για μια εργασία, πρέπει όλες να συμπεριλαμβάνονται σε ένα μοναδικό Requirements attribute, με χρήση μιας boolean έκφρασης.

Για παράδειγμα, ας θεωρήσουμε ότι ο χρήστης επιθυμεί να εκτελέσει την εργασία του σε ένα CE που χρησιμοποιεί το PBS για batch system, και του οποίου τα WNs έχουν τουλάχιστον δύο CPUs. Η αντίστοιχη απαίτηση στο JDL αρχείο θα εκφραζόταν με το ακόλουθο attribute:

```
Requirements = other.GlueCEInfoLRMSType == "PBS" &&  
other.GlueCEInfoTotalCPUs > 1;
```

Η ανάθεση μιας εργασίας σε ένα συγκεκριμένο CE μπορεί να επιτευχθεί π.χ. με την ακόλουθη έκφραση:

```
Requirements = other.GlueCEUniqueID ==  
"lxshare0286.cern.ch:2119/jobmanager-pbs-short";
```

Η έκφραση ενός requirement μπορεί να γίνει και με ένα regular expression. Έστω για παράδειγμα ότι ένας χρήστης επιθυμεί όλες οι εργασίες του να εκτελεστούν σε οποιοδήποτε CE που ανήκει στο domain cern.ch. Αυτό μπορεί να επιτευχθεί με την ακόλουθη έκφραση:

```
Requirements = RegExp("cern.ch", other.GlueCEUniqueID);
```

Η ακριβώς αντίθετη απαίτηση εκφράζεται ως εξής:

```
Requirements = (!RegExp("cern.ch", other.GlueCEUniqueID));
```

Ένα πιο σύνθετο παράδειγμα είναι το ακόλουθο:

```
Requirements = other.GlueHostNetworkAdapterOutboundIP==true &&  
Member("VO-alice-  
AliEn", other.GlueHostApplicationSoftwareRunTimeEnvironment) &&  
Member("VO-alice-ALICE-v4-  
01", other.GlueHostApplicationSoftwareRunTimeEnvironment) &&  
(other.GlueCEPolicyMaxWallClockTime > 86000) ;
```

Στην παραπάνω έκφραση, επιλέγεται το CE που έχει εγκατεστημένα συγκεκριμένα πακέτα λογισμικού (VO-alice-AliEn and VO-alice-ALICE-v4-01-Rev-01) και επιτρέπει στην εργασία να εκτελεστεί για παραπάνω από 86000 δευτερόλεπτα (έτσι ώστε να μην απορριφθεί η εργασία προτού προλάβει να τελειώσει επιτυχώς).

#### **8.4.2.2 Rank**

Η επιλογή του CE στο οποίο θα εκτελεστεί μια εργασία ανάμεσα σε αυτά που ικανοποιούν τις απαιτήσεις που εκφράζονται στο Requirements attribute γίνεται με βάση το rank του CE, το οποίο είναι μια ποσότητα που ορίζεται ως ένας floating-point αριθμός. Το CE που τελικά επιλέγεται είναι αυτό με το μεγαλύτερο rank. Εξ ορισμού, η παράμετρος Rank υπολογίζεται ως `-other.GlueCEStateEstimatedResponseTime`, όπου ο εκτιμώμενος χρόνος απόκρισης είναι μια προσέγγιση του χρονικού διαστήματος μεταξύ της υποβολής της εργασίας και την έναρξη της εκτέλεσης της εργασίας. Ο χρήστης έχει τη δυνατότητα να επανακαθορίσει το rank κάνοντας χρήση του Rank attribute στο JDL αρχείο και χρησιμοποιώντας εκφράσεις που σχετίζονται με τα χαρακτηριστικά του CE. Για παράδειγμα, το attribute:

```
Rank = other.GlueCEStateFreeCPUs
```

θα κατατάξει ως καλύτερο το CE με τους περισσότερους ελεύθερους CPUs.

Από την άλλη, η ακόλουθη έκφραση:

```
Rank = ( other.GlueCEStateWaitingJobs == 0 ? other.GlueCEStateFreeCPUs :  
-other.GlueCEStateWaitingJobs );
```

ορίζει ότι το CE που θα επιλεγεί θα είναι αυτό με τις λιγότερες εργασίες σε αναμονή, ή αυτό με τους περισσότερους ελεύθερους CPUs, αν δεν υπάρχουν καθόλου εργασίες σε αναμονή.

Επιπρόσθετα υπάρχει η δυνατότητα προστασίας από τα λάθη. Αυτό επιτυγχάνεται με την παράμετρο `RetryCount` στο αρχείο JDL. Η παράμετρος αυτή καθορίζει πόσες φορές θα

επαναληφθεί η εκτέλεση της εργασίας στον ίδιο ή άλλο υπολογιστικό πόρο. Η δυνατότητα αυτή είναι ιδιαίτερα σημαντική, δεδομένης της σημασίας εκτέλεσης εργασιών υπολογιστικής ογκολογίας για τους χρήστες.

## **8.5 Λεπτομέρειες υλοποίησης**

Η υλοποίηση της εφαρμογής πραγματοποιήθηκε με τρόπο που εξασφαλίζει:

- *Ευχρηστία*, μέσα από το φιλικό περιβάλλον του φυλλομετρητή (Web Browser) χωρίς την ανάγκη για εξεζητημένες γνώσεις σχετικά με κατακευματισμένα υπολογιστικά συστήματα και πλέγματα.
- *Αποδοτικότητα*, με την γραφική απεικόνιση των υπηρεσιών πλέγματος στον χρήστη.
- *Ανεξαρτησία από λειτουργικό σύστημα*, με την χρήση μόνο ενός φυλλομετρητή που είναι διαθέσιμος σε κάθε πλατφόρμα, καθώς και την
- *μη εγκατάσταση λογισμικού* στους χρήστες και *κεντρική διαχείριση*, με την εγκατάσταση της πύλης σε έναν κεντρικό εξυπηρετητή και της απομακρυσμένης πρόσβασης των χρηστών και διαχειριστών σε αυτόν και τις υπηρεσίες του.

Η εφαρμογή επίσης θα έχει την δυνατότητα ενσωμάτωσης προηγμένων υπηρεσιών Grid υψηλού επιπέδου όπως:

- Υιοθέτηση και υλοποίηση πολιτικών που θα υλοποιούν Ποιότητα Υπηρεσίας (QoS) για τις εφαρμογές των χρηστών που εγκαθίστανται στο σύστημα.
- Προηγμένο σύστημα διαμεσολάβησης και δέσμευσης πόρων (brokers) και βέλτιστη αξιοποίησή τους, λαμβάνοντας υπόψη ειδικές ανάγκες εφαρμογών, όσων αφορά στη χρήση της φύσης των πόρων.

## 8.6 Καινοτόμος αρχιτεκτονική προσέγγιση

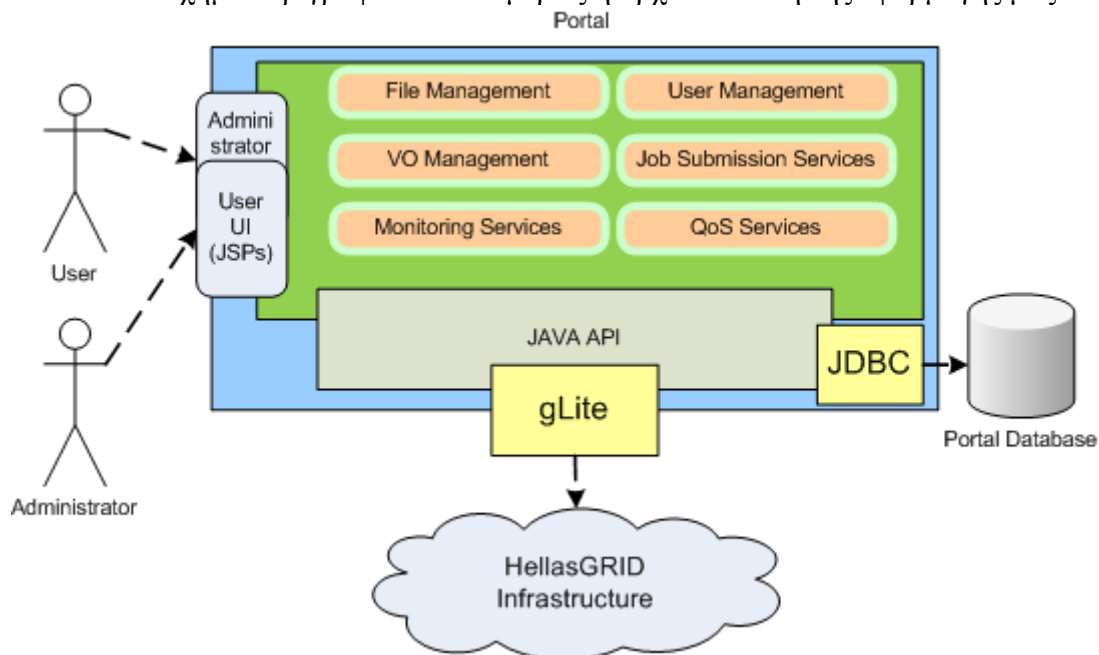
Για το ν σχεδιασμό και την ανάπτυξη της δικτυακής πύλης και της εφαρμογής ακολουθήθηκε πολύ-επίπεδη αρχιτεκτονική (multi-tier) [48]. Η συγκεκριμένη αρχιτεκτονική προσέγγιση εξασφαλίζει διαστρωμάτωση στην λειτουργικότητα και αποδοτικότητα της εφαρμογής τόσο κατά τον σχεδιασμό της όσο και κατά την εγκατάσταση και λειτουργία της. Ειδικότερα σε πολύπλοκες εφαρμογές που συνδυάζουν enterprise χαρακτηριστικά με την παρουσίαση και διάθεση τους στον τελικό χρήστη, η συγκεκριμένη προσέγγιση καθίσταται απαραίτητη.

Αναλυτικά τα πλεονεκτήματά της είναι:

- **Ευκολότερη ανάπτυξη:** Η οικοδόμηση των τριών επιπέδων εφαρμογών είναι ευκολότερη δεδομένου ότι τα συστατικά μπορούν να επαναχρησιμοποιηθούν από περισσότερες από μία φορές.
- **Ευκολότερη συντήρησης:** Η λειτουργικότητα του κάθε επιπέδου είναι σαφής με τα τυχόν προβλήματα να περιλαμβάνονται σε μια θέση μόνο και έτσι ο εντοπισμός και η αντιμετώπισή τους γίνεται πιο άμεσα και μεθοδευμένα.
- **Τυποποίηση της εφαρμογής:** Οι επιχειρησιακές λειτουργίες είναι τοποθετημένες σε ένα επίπεδο ενώ η πρόσβαση σε αυτές μπορεί να είναι από πολλαπλάσιες θέσεις.
- **Περισσότερη ευελιξία:** Οι πολύ-επίπεδες εφαρμογές δίνουν στους υπεύθυνους ανάπτυξης τη δυνατότητα να συνδυάσουν στην επιχειρησιακή λειτουργία με την παρουσίαση καθώς και να την επέκτασή τους με επιπλέον δυνατότητες που δεν περιλαμβανόταν στον αρχικό σχεδιασμό.

- **Καλύτερη ποιότητα:** Επειδή η εφαρμογή χτίζεται σε διακριτικές μονάδες-επίπεδα, αυτές οι μονάδες μπορούν να εξεταστούν λεπτομερώς και να γίνουν περισσότερο στιβαρές κατά τη διάρκεια του χρόνου.
- **Ευκολότερη συνεργασία διαφορετικού προσωπικού ανάπτυξης:** Η multi-tier αρχιτεκτονική δίνει την ευελιξία κάθε εμπλεκόμενος στην ανάπτυξη της εφαρμογής να ασχολείται με συγκεκριμένο αντικείμενο και να το υλοποιήσει χρησιμοποιώντας τα διαφορετικά εργαλεία.
- **Επιτρέπει τη χρήση των νέων εργαλείων ανάπτυξης και τεχνολογιών:** Είναι ευκολότερο να εισαχθούν οι τεχνολογίες σε μια εφαρμογή που χωρίζεται στα πολλαπλάσια στρώματα.

Στο ακόλουθο σχήμα περιγράφεται λεπτομερώς η αρχιτεκτονική της εφαρμογής μας:



Σχήμα 41: Αναπαράσταση της υλοποίησης και λειτουργίας της εφαρμογής

Τα επίπεδα υλοποίησης της συγκεκριμένης εφαρμογής είναι:

- Παρουσίαση (Interface): Administrator and User UIs (JSPs)

*Η σελίδα αυτή είναι σκόπιμα λευκή*



- Λειτουργικότητας (Business): File Management, User Management, VO Management, Job Submission Services, Monitoring Services and QoS Services
- Βάσης Δεδομένων (Database)
- Πλέγματος (Grid)

### **8.6.1 Παρουσίαση**

Το επίπεδο αυτό περιλαμβάνει τις λειτουργίες αλληλεπίδρασης με τον χρήστη. Είναι υλοποιημένη με τέτοιο τρόπο που καθιστά την επικοινωνία του χρήστη με το πλέγμα απλή, εύκολη και αποδοτική. Εδώ όλες οι δυνατότητες που έχουν υλοποιηθεί στην εφαρμογή προσφέρονται στους χρήστες μέσα από JSP σελίδες. Ειδικότερα, το επίπεδο παρουσίασης δίνει την δυνατότητα στους χρήστες να πραγματοποιούν όλες τις απαραίτητες ενέργειες για υποβολή εργασιών υπολογιστικής βιολογίας στο πλέγμα, με υψηλό επίπεδο ποιότητας υπηρεσιών χωρίς την ανάγκη γνώσεων σχετικών με τα πολύπλοκα υπολογιστικά συστήματα που θα τις εκτελέσουν. Όλες οι λειτουργίες παρουσιάζονται φιλικά μέσα από ένα τελευταίας τεχνολογίας γραφικό περιβάλλον. Πρέπει να σημειωθεί ότι το επίπεδο παρουσίασης υποστηρίζει περισσότερες από μία κατηγορίες χρηστών. Έτσι λειτουργικότητα που έχουν πρόσβαση οι απλοί χρήστες είναι περιορισμένη σε σχέση με τους διαχειριστές (administrators) της εφαρμογής.

### **8.6.2 Λειτουργικότητα**

Στο επίπεδο αυτό υλοποιούνται όλες οι λειτουργίες της εφαρμογής και επιτυγχάνεται η διασύνδεση μεταξύ του επιπέδου παρουσίασης και των επιπέδων βάσης δεδομένων και πλέγματος. Για το την υλοποίηση του συγκεκριμένου στρώματος χρησιμοποιήθηκε η γλώσσα προγραμματισμού JAVA λόγω των αυξημένων δυνατοτήτων της και της αποδοτικής διασύνδεσης με τα άλλα στρώματα της εφαρμογής. Εδώ αναπτύχθηκαν οι λειτουργίες της εφαρμογής για: Διαχείριση Αρχείων, Χρηστών και Πόρων πλέγματος

όπως επίσης και οι υπηρεσίες υποβολής εργασιών στο πλέγμα και επίβλεψής τους. Παράλληλα, κάθε μία από τις επιμέρους λειτουργίες αλλά και συνολικά η εφαρμογή ενισχύθηκε με δυνατότητες παροχής υπηρεσιών υψηλού επιπέδου (QoS) με στόχο την αποδοτικότερη αξιοποίηση των υπολογιστικών πόρων και τη μείωση του χρόνου απόκρισης για της υποβληθείσες εργασίες.

### **8.6.3 Βάση Δεδομένων**

Η βάση δεδομένων λειτουργεί συμπληρωματικά στο επίπεδο λειτουργικότητας της εφαρμογής. Εδώ διατηρούνται όλα τα απαραίτητα δεδομένα για την διαχείριση της εφαρμογής ενώ ταυτόχρονα αποθηκεύονται QoS πληροφορίες για τους πόρους της υποδομής πλέματος. Οι πληροφορίες αυτές αξιοποιούνται στη συνέχεια για την αποδοτικότερη δρομολόγηση των νέων εργασιών που θα εκτελεστούν στο πλέγμα.

### **8.6.4 Πλέγμα**

Στο επίπεδο αυτό περιλαμβάνονται όλες οι λειτουργίες επικοινωνίας με το πλέγμα, τους πόρους και τις υπηρεσίες του. Ειδικότερα εδώ πραγματοποιείται η διασύνδεση της εφαρμογής με το HellasGrid [98][71][93][69] μέσω του λογισμικού gLite που έχει εγκατασταθεί στον ίδιο εξυπηρετητή με την εφαρμογή.

## **8.7 Αποτελέσματα**

### **8.7.1 Σενάριο χρήσης: Ακτινοθεραπεία**

Ο κώδικας προσομοίωσης της συγκεκριμένης θεραπευτικής μεθόδου [86][29][79] περιέχει ένα σχετικά μεγάλο αριθμό παραμέτρων. Σε πρώτη φάση έγινε επιλογή ενός υποσυνόλου των μεταβλητών αυτών, ώστε να γίνει δυνατή η αυτόματη εκτέλεση παραμετρικών προσομοιώσεων στο πλέγμα ως προς αυτό το υποσύνολο και ώστε να μελετηθεί η συμπεριφορά της συγκεκριμένης θεραπευτικής μεθόδου για ένα μεγάλο εύρος

των μεταβλητών εισόδου. Οι παράμετροι που επιλέχθηκαν για τις παραμετρικές προσομοιώσεις είναι οι εξής:

```
const unsigned short SelectedFractionationScheme = 2;  
const float dose = 1.2;  
const double alpha_P = 0.17;  
const double beta_P = 0.02;  
const long newsize = 96;  
const unsigned long TC = 40;  
const string image = "Input96.raw";
```

Οι πρώτες έξι παράμετροι σχετίζονται με το θεραπευτικό σχήμα που θα προσομοιωθεί, ενώ η τελευταία παράμετρος είναι ένα αρχείο με απεικονιστικά δεδομένα της σχετικής κατάστασης του ασθενή.

Για την εκτέλεση μιας παραμετρικής προσομοίωσης απαιτείται από τον τελικό χρήστη να παρέχει ένα σύνολο συνδυασμών των παραπάνω παραμέτρων. Για παράδειγμα, για μελέτη των αποτελεσμάτων διαφορετικών θεραπευτικών σχημάτων ακτινοθεραπείας στον ίδιο ασθενή, δίνεται ένα σύνολο συνδυασμών των πρώτων έξι παραμέτρων, ενώ τα δεδομένα του ασθενή (παράμετρος “image”) παραμένουν σταθερά. Αντίστροφα, διατηρώντας σταθερές τις έξι πρώτες παραμέτρους και δίνοντας διαφορετικά απεικονιστικά δεδομένα, μπορεί να μελετηθεί η συμπεριφορά ενός συγκεκριμένου θεραπευτικού σχήματος για διαφορετικούς ασθενείς.

Στα πρώτα στάδια προσαρμογής της εφαρμογής της ακτινοθεραπείας, πραγματοποιήθηκε μια σειρά από παραμετρικές προσομοιώσεις στο πλέγμα [97][22][30][103]. Συγκεκριμένα, έγινε προσομοίωση για 6 ακτινοθεραπευτικά σχήματα, με 11 τιμές κυτταρικού κύκλου TC και 7 ζεύγη τιμών alpha\_P, beta\_P. Η προσομοίωση για κάθε ένα από τα θεραπευτικά αυτά σχήματα αποτελείται δηλαδή από 77 συνδυασμούς παραμέτρων. Κάθε θεραπευτικό σχήμα αντιμετωπίστηκε ως μια παραμετρική προσομοίωση και υποβλήθηκε για εκτέλεση στο πλέγμα. Αρχικά, σε κάθε εργασία που

υποβλήθηκε στο πλέγμα ανατέθηκε η εκτέλεση της προσομοίωσης για ένα σετ παραμέτρων (συνολικά δηλαδή υποβλήθηκαν 77 εργασίες για κάθε θεραπευτικό σχήμα).

Στον Πίνακα 1 παρατίθενται κάποια ενδεικτικά στατιστικά στοιχεία σχετικά με το χρόνο εκτέλεσης των εργασιών στο πλέγμα για τα 6 διαφορετικά θεραπευτικά σχήματα.

<b>Σχήμα #</b>	<b>Μέσος Χρόνος Εκτέλεσης Εργασίας</b>	<b>Συνολικός Χρόνος Εκτέλεσης Σχήματος</b>
<b>1</b>	~32 mins	~58 mins
<b>2</b>	~31 mins	~59 mins
<b>3</b>	~34 mins	~59 mins
<b>4</b>	~38 mins	~72 mins
<b>5</b>	~36 mins	~66 mins
<b>6</b>	~21 mins	~47 mins

**Πίνακας 6: Στατιστικά χρόνων εκτέλεσης στο πλέγμα**

Ο μέσος χρόνος εκτέλεσης εργασίας είναι ο μέσος καθαρός χρόνος που απαιτείται για την εκτέλεση της προσομοίωσης στο WN του πλέγματος που αυτή ανατέθηκε. Ο συνολικός χρόνος εκτέλεσης σχήματος είναι το χρονικό διάστημα που μεσολαβεί από τη στιγμή που υποβάλλεται η πρώτη από τις 77 εργασίες στο πλέγμα μέχρι να ληφθούν τα αποτελέσματα και από την τελευταία από αυτές. Παρατηρούμε ότι η χρήση του πλέγματος επιτυγχάνει σημαντική επιτάχυνση στο συνολικό χρόνο εκτέλεσης που απαιτείται για την ολοκλήρωση της εκτέλεσης κάθε σετ προσομοιώσεων. Προφανώς, το όφελος από την επιτάχυνση της εκτέλεσης των προσομοιώσεων που επιτυγχάνεται θα είναι σημαντικότερο όσο αυξάνει ο αριθμός των θεραπευτικών σχημάτων προς αξιολόγηση και όσο αυξάνεται ο αριθμός των ασθενών.

Η προσέγγιση που περιγράφηκε προηγουμένως για ανάθεση μιας προσομοίωσης ανά υποβαλλόμενη εργασία έχει ορισμένα μειονεκτήματα και πλεονεκτήματα. Καταρχήν, το προφανές πλεονέκτημα είναι ότι με αυτόν τον τρόπο αυξάνεται ο βαθμός

παραλληλοποίησης στην εκτέλεση της παραμετρικής προσομοίωσης, αφού δυνητικά μπορούν να εκτελούνται όλες οι εργασίες ταυτόχρονα, επιτυγχάνοντας με αυτό τον τρόπο το μέγιστο δυνατό speedup χωρίς να προχωρήσουμε σε παραλληλοποίηση σε επίπεδο κώδικα προσομοίωσης. Ένα μειονέκτημα είναι ότι αυτό το μέγιστο speedup μπορεί να επιτευχθεί μόνο σε περίπτωση που όλες οι εργασίες ανατίθενται σε ελεύθερους κόμβους χωρίς καθυστέρηση σε ουρές αναμονής κτλ. Επιπλέον, η υποβολή μεγάλου αριθμού εργασιών αυξάνει το συνολικό overhead στο πλέγμα λόγω πολλαπλών data staging. Τέλος, λαμβάνοντας υπόψη ότι η υποβολή των εργασιών πρέπει τελικά να γίνεται από την πύλη που αναπτύχθηκε για τις ανάγκες του έργου, το σημαντικότερο μειονέκτημα της συγκεκριμένης προσέγγισης είναι ο αυξημένος αριθμός εργασιών προκαλεί μεγάλο υπολογιστικό φόρτο στην πύλη λόγω υποβολής, επίβλεψης και ανάκτησης των αποτελεσμάτων των εργασιών. Το φορτίο μπορεί να γίνει εξαιρετικά μεγάλο όσο αυξάνεται ο αριθμός των χρηστών που χρησιμοποιούν την πύλη.

### **8.7.2 Σενάριο χρήσης: Χημειοθεραπεία**

Η δεύτερη θεραπευτική μέθοδος υπολογιστικής ογκολογίας που μεταφέρθηκε στο περιβάλλον του πλέγματος είναι η μέθοδος της χημειοθεραπείας. Για τους ίδιους λόγους που περιγράφηκαν και για την ακτινοθεραπεία, επιλέχτηκε η προσέγγιση εκτέλεσης πολλαπλών προσομοιώσεων ανά υποβαλλόμενη εργασία. Προς το παρόν δεν έχει πραγματοποιηθεί κάποια αντίστοιχη παραμετρική μελέτη, όπως στην περίπτωση της ακτινοθεραπείας, παρά μόνο κάποιες δοκιμαστικές εκτελέσεις. Τα αποτελέσματα αναφορικά με την απόδοση της εφαρμογής στο πλέγμα αναμένεται να είναι αντίστοιχα με αυτά που προέκυψαν από τις παραμετρικές προσομοιώσεις της ακτινοθεραπείας. Στο άμεσο μέλλον θα γίνει εκμετάλλευση της πλεγματικής υποδομής για παρόμοιες παραμετρικές προσομοιώσεις για τη μέθοδο της χημειοθεραπείας.

## **8.8 Συμπεράσματα και μελλοντικά σχέδια**

Από τα πρώτα στάδια ανάπτυξης της εφαρμογής, αυτή ήταν διαθέσιμη στους τελικούς χρήστες, ώστε να συλλέξουμε σχόλια που θα οδηγούσαν στην βελτίωσή και αποσφαλμάτωσή της. Η διαδικασία αυτή επαναλήφθηκε αρκετές φορές μέχρι την τελική έκδοση της εφαρμογής. Τα αποτελέσματα μέχρι στιγμής είναι πολύ θετικά και οι χρήστες είχαν την δυνατότητα να εκτελέσουν δεκάδες εξομοιώσεις με εύκολο και αποδοτικό τρόπο.

Στο μέλλον, προγραμματίζουμε την επέκταση της εφαρμογή προσομοίωσης προκειμένου να εξεταστούν στο μέγιστο οι απαιτήσεις χρηστών για την ποιότητα της υπηρεσίας (QoS). Οι νέοι μηχανισμοί είναι υπό ανάπτυξη για να παρέχουν τον προηγμένη δρομολόγηση εργασίας βασισμένη

- στις ανάγκες των χρηστών,
- τις ικανότητες των πόρων πλέγματος και
- την εμπειρία των προηγούμενων εκτελέσεων εργασίας.

Οι μηχανισμοί αυτοί θα περιλαμβάνουν επιπρόσθετες λειτουργίες και δυνατότητες δρομολόγησης με την χρήση πιο σύνθετων αλγορίθμων.

Παράλληλα θα διορθωθούν τυχών αδυναμίες της εφαρμογής σε θέματα ασφαλείας. Τέλος η εφαρμογή θα γίνει διαθέσιμη σε ένα ευρύτερο κύκλο χρηστών έτσι ώστε να αξιοποιηθεί όσο το δυνατό περισσότερο από την ιατρική και ερευνητική κοινότητα και τα αποτελέσματα αυτά να οδηγήσουν στην επιπλέον βελτίωση και αναβάθμισή της.

# 9

## **Συμπεράσματα και μελλοντικές εργασίες**

Στο κεφάλαιο αυτό συνοψίζουμε τα αποτελέσματα και την προσφορά της διατριβής στον τεχνολογικό τομέα των υπηρεσιοστρεφών υποδομών και αρχιτεκτονικών. Επιπλέον γίνεται αναφορά σε μερικά από τα θέματα μελλοντικής εργασίας και επέκτασης των ερευνητικών αποτελεσμάτων της διατριβής.

### **9.1 Συμπεράσματα**

Το κίνητρο για τη ερευνητική εργασία στα πλαίσια της διατριβής ήταν η μετάβαση από τις ερευνητικές υποδομές πλέγματος στις επιχειρηματικές και η αυξανόμενη ανάγκη των χρηστών για ποιότητα υπηρεσιών και διασφάλισή της μέσα σε αυτές. Γρήγορα διαπιστώσαμε πως ο συγκεκριμένος ερευνητικός τομέας είναι έντονα ευμετάβλητος κάτι που αποτυπωνόταν και στις αρχιτεκτονικές των υποδομών. Ο λόγος για αυτό ήταν η διαφοροποίηση των επιχειρηματικών μοντέλων από τη μία εφαρμογή στην άλλη καθώς ακόμα και στην ίδια εφαρμογή ανάλογα με τις απαιτήσεις των χρηστών και των τρόπο που επιθυμούν να χρησιμοποιήσουν τις υπηρεσίες της εφαρμογής.

Κοινό χαρακτηριστικό σε όλα τα επιχειρηματικά μοντέλα είναι η ανάγκη για διασφάλιση της ποιότητας υπηρεσιών ανάλογα με τις απαιτήσεις των χρηστών. Η διατριβή παρακολούθησε τις αλλαγές και προσπάθησε να καλύψει ακριβώς αυτή την ανάγκη μέσα

από καινοτόμες αρχιτεκτονικές λύσεις και μηχανισμούς εφαρμόσιμους σε κάθε σύγχρονη υπηρεσιοστρεφή υποδομή όπως τα πλέγματα υπολογιστών, υποδομές SaaS (Software as a Service), Cloud Computing αλλά και νέες υποδομές που ενδεχομένως προκύψουν στο μέλλον. Για να το πετύχουμε αυτό, αρχικά πραγματοποιήσαμε ανάλυση των επιχειρηματικών μοντέλων που σχετίζονται με υπηρεσιοστρεφείς υποδομές. Στα πλαίσια αυτής της ανάλυσης καταλήξαμε σε συγκεκριμένες αρχές αρχιτεκτονικού σχεδιασμού τις οποίες στη συνέχεια εφαρμόσαμε σε εφαρμογές βιοϊατρικής, πολυμέσων και πραγματικού χρόνου απόκρισης.

Συγκεκριμένα παρουσιάσαμε καινοτόμους μηχανισμούς τόσο από την πλευρά του καταναλωτή της υπηρεσίας για την αποδοτική επιλογή παρόχων όσο και από την πλευρά των παρόχων για αποδοτική διαχείριση των πόρων και υπηρεσιών με τη χρήση ευέλικτων και προσαρμόσιμων ανάλογα με την εφαρμογή πολιτικών. Οι συμφωνίες επιπέδου υπηρεσιών έχουν κεντρικό ρόλο στο σχεδιασμό και τη λειτουργία των προτεινόμενων υποδομών. Με τη χρήση συμφωνιών επιπέδου υπηρεσιών διασφαλίζεται ο τρόπος έκφρασης των παραμέτρων ποιότητας υπηρεσιών με μια κοινή γλώσσα, αντιληπτή και από τις δύο πλευρές, χρήστη και παρόχου, ενώ πάνω σε αυτές τις συμφωνίες στηρίζεται η επιχειρηματική τους σχέση. Επιπρόσθετα δίνεται η δυνατότητα έκφρασης επιπλέον πληροφοριών επιχειρηματικού επιπέδου όπως αποζημιώσεις σε περίπτωση αποτυχίας, αλλά και διαχειριστικού επιπέδου με πολιτικές χειρισμού των γεγονότων που επηρεάζουν την ποιότητα υπηρεσιών. Με αυτό τον τρόπο κάθε πλευρά μπορεί και εκφράζει τις επιθυμίες της με τρόπο που η ίδια αντιλαμβάνεται και εξαλείφεται η ανάγκη από την πλευρά των χρηστών για γνώση της υποδομής και των παραμέτρων των πόρων και από την πλευρά των παρόχων για τη κατανόηση της εφαρμογής.

Οι αρχιτεκτονικές που παρουσιάστηκαν για τις υπηρεσιοστρεφείς υποδομές είναι επιπλέον ανοιχτές και προσαρμόσιμες. Με αυτόν τον τρόπο, ανάλογα με τον τύπο της



κάθε εφαρμογής, η υποδομή μπορεί να προσαρμοστεί εμπλέκοντας περισσότερα ή λιγότερα λειτουργικά συστατικά ή υπηρεσίες. Τόσο η γενική αρχιτεκτονική όσο και οι υπηρεσίες της είναι εύκολο να επεκταθούν ώστε να καλύπτουν τις εξειδικευμένες ενδεχομένως ανάγκες της εκάστοτε εφαρμογής. Χαρακτηριστικό παράδειγμα είναι ο μηχανισμός διαχείρισης με τη χρήση πολιτικών ο οποίος ανάλογα με την εφαρμογή αλλά και τις απαιτήσεις των χρηστών μπορεί να χρησιμοποιεί διαφορετικές πηγές πληροφόρησης αλλά και αλγόριθμους λήψης αποφάσεων, αυξάνοντας της αυτονομία της και την αποδοτικότητα της εφαρμογής.

## **9.2 Μελλοντικές εργασίες**

Οι υπηρεσιοστρεφείς αρχιτεκτονικές και υποδομές αποτελούν την τάση σήμερα για τις τεχνολογίες πληροφορικής και επικοινωνιών. Παρά το γεγονός ότι αναπτύσσονται και υλοποιούνται νέες μέθοδοι και τεχνολογίες για τα συστήματα υπολογιστών και επικοινωνιών, ο σχεδιασμός τους στην πλειοψηφία των περιπτώσεων ακολουθεί υπηρεσιοστρεφείς αρχές. Χαρακτηριστικό παράδειγμα είναι οι τεχνολογίες Cloud Computing που ουσιαστικά βασίζονται στις προϋπάρχουσες τεχνολογίες πλέγματος με την προσθήκη μηχανισμών και μεθόδων εικονοποίησης. Ιδιαίτερο ενδιαφέρον παρουσιάζει λοιπόν η μελλοντική αξιοποίηση των αποτελεσμάτων της διατριβής σε υποδομές Cloud Computing και σε virtualized πόρους και υπηρεσίες. Η έρευνα αναφορικά με τα επιχειρηματικά μοντέλα και την ποιότητα υπηρεσιών θα πρέπει να συνεχιστεί και στις νέες υπηρεσιοστρεφείς υποδομές ώστε να διασφαλιστεί με αυτό τον τρόπο η μελλοντική λειτουργικότητα και η οικονομική βιωσιμότητά τους.

*Η σελίδα αυτή είναι σκόπιμα λευκή*

# 10

## Συντομογραφίες

Στον παρακάτω πίνακα παραθέτουμε τους όρους και συντομογραφίες που χρησιμοποιήθηκαν στη διατριβή:

<b>Όρος</b>	<b>Ερμηνεία</b>
<b>CORBA</b>	Common Object Request Broker Architecture
<b>EGEE</b>	Enabling Grids for E-science
<b>GGF</b>	Global Grid Forum
<b>GUI</b>	Graphical User Interface
<b>ICT</b>	Information and Communication Technology
<b>IRMOS</b>	Interactive Real-time Multimedia Applications on Service Oriented Infrastructures
<b>ISONI</b>	Intelligent Service Oriented Network Infrastructure
<b>NextGRID</b>	Next Generation Grid Architecture Integrated Project
<b>OASIS</b>	Organization for the Advancement of Structured Information Standards
<b>OGSA</b>	Open Grid Services Architecture
<b>QoS</b>	Quality of Service
<b>SaaS</b>	Software as a Service
<b>SLA</b>	Service Level Agreement
<b>SOA</b>	Service Oriented Architecture
<b>SOAP</b>	Simple Object Access Protocol
<b>SOI</b>	Service Oriented Infrastructure
<b>WSDL</b>	Web Services Description Language
<b>WSRF</b>	Web Services Resource Framework

*Η σελίδα αυτή είναι σκόπιμα λευκή*

# 11

## *Βιβλιογραφικές Αναφορές*

- [1] A. Andrieux, K. Czajkowski, A. Dan, K. Keahey, H. Ludwig, J. Pruyne, J. Rofrano, S. Tuecke, and M. Xu, “Web Services Agreement Specification (WS-Agreement),” Global Grid Forum, May 2004.
- [2] A. Keller, H. Ludwig, The WSLA framework: specifying and monitoring service level agreements for Web Services, *J. Network Systems Manag.* 11 (1) (2003).
- [3] A. Menychtas, D. Apostolopoulos, D. Kyriazis, K. Christodouloupoulos, H. Avramopoulos and Theodora Varvarigou, “Enabling a Network Simulation Application on Grid Infrastructure”, 11th Panhellenic Conference on Informatics, Patras, Greece, 2007.
- [4] Al-ali, R. et al., 2004. “Analysis and provision of QoS for distributed grid applications”, *J. Grid Comput.* 2, 2, 163-182.
- [5] Alan O. Freier, Philip Karlton, and Paul C. Kocher. The SSL protocol: Version 3.0.
- [6] Aloisio G., Cafaro M., Fiore S., et al., “ProGenGrid: A Workflow Service Infrastructure for Composing and Executing Bioinformatics Grid Services”, Proceedings of the 18th IEEE Symposium on Computer-Based Medical Systems, Dublin, Ireland, pp.555-560, 2005.
- [7] Amazon Elastic Compute Cloud (Amazon EC2), <http://aws.amazon.com/ec2/>
- [8] Andreas Menychtas, Dimosthenis Kyriazis, Konstantinos Tserpes, “Real-time Reconfiguration for Guaranteeing QoS Provisioning Levels in Grid Environments”, Special Issue on Real-time attributes in Grids, Future Generation Computer Systems, 2008.

- [9] Andreas Menychtas, Konstantinos Dolkas, Dimosthenis Kyriazis and Theodora Varvarigou, “Building Portals to access Grid Middleware”, GGF 14 Conference Chicago – USA, 27-30 June 2005.
- [10] Andronico A., Barbera R., Falzone A., et al., “GENIUS: a simple and easy way to access computational and data grids”, *Future Generation Computer Systems*, Volume 19, Issue 6, pp. 805-813, 2003.
- [11] Angulo D., Foster I., Liu C., et al., "Design and Evaluation of a Resource Selection Framework for Grid Applications", *Proceedings of the 11th IEEE International Symposium on High Performance Distributed Computing (HPDC-11)*, Edinburgh, Scotland, pp. 63-73, 2002.
- [12] Antonios Litke, Kleopatra Konstanteli, Vassiliki Andronikou, Sotirios Chatzis, Theodora Varvarigou, *Managing service level agreement contracts in OGSA-based Grids*, *Future Generation Computer Systems*, Volume 24, Issue 4, April 2008.
- [13] Atkinson, B., et al.: *Web Services Security (WS-Security)*. Technical report, Microsoft, IBM and Verisign, April 2002.
- [14] Beeson B., Melnikoff S., Venugopal S., et al., “A Portal for Grid-enabled Physics”, *Proceedings of the 2005 Australasian workshop on Grid computing and e-research*, pp. 13-20, 2005.
- [15] Bellwood, T.; Clément, L.; Ehnebuske, D.; Hately, A.; Hondo, M.; Husband, Y.L.; Januszewski, K.; Lee, S.; McKee, B.; Munter, J.; and von Riegen, C. 2002. *Universal Description Discovery & Integration (UDDI) Specification*.
- [16] Buyya R., Abramson D., Giddy J., “Economy Driven Resource Management Architecture for Computational Power Grids”, *Proceedings of the 2000 International Conference on Parallel and Distributed Processing Techniques and Applications*, Las Vegas, USA, 2000.
- [17] Cardoso J., Sheth A., Miller J., et al. , “Quality of Service for Workflows and Web Service Processes”, *Journal of Web Semantics*, Volume 1, Issue 3, pp. 281-308, 2004.
- [18] Casanova H., Obertelli G., Berman F., et al., “The AppLeS Parameter Sweep Template: User-level middleware for the Grid”, *Journal of Scientific Programming*, Volume 8, Issue 3, pp. 111-126, 2000.

- [19] D. Kyriazis, K. Tserpes, A. Menychtas, A. Litke and T. Varvarigou, “An innovative Workflow Mapping Mechanism for Grids in the frame of Quality of Service”, on Future Generation Computer Systems, 2008.
- [20] D. Kyriazis, K. Tserpes, A. Menychtas, I. Sarantidis and T. Varvarigou, “Service Selection and Workflow Mapping for Grids: An approach exploiting Quality of Service Information”, on Concurrency and Computation: Practice and Experience, Willey Interscience, 2008.
- [21] Debusmann, M.; Keller, A., "SLA-driven management of distributed systems using the common information model," Integrated Network Management, 2003. IFIP/IEEE Eighth International Symposium on, vol., no., pp. 563-576, 24-28 March 2003.
- [22] Deiscboek TS, Zhang L, Yoon J et al. “In silico cancer modeling: is it ready for prime time”, Nat Clin Pra Oncol 6(1): 34-42, 2009.
- [23] Dimitra Dionysiou, Georgios Stamatakos, Theodoros Athanaileas, Andreas Menychtas, Dimitra Kaklamani, Theodora Varvarigou and Nikolaos Uzunoglu, “In Silico Simulation of a Clinical Trial Concerning Tumour Response to Radiotherapy”, e-conference IECCS 2007.
- [24] Dimosthenis Kyriazis, Andreas Menychtas and Theodora Varvarigou, “Grid Workflows with encompassed Business Relationships: An approach establishing Quality of Service Guarantees” in book “Quantitative Quality of Service for Grid Computing”, IGI Global Publishers, 2008.
- [25] Dimosthenis Kyriazis, Andreas Menychtas, Dimitra Dionysiou, Georgios Stamatakos, Theodora Varvarigou, “Clinical Trial Simulation in Grid Environments”, 8th International Conference on BioInformatics and BioEngineering (BIBE), Athens, Greece, 2008.
- [26] Dimosthenis Kyriazis, Andreas Menychtas, Konstantinos Tserpes, Theodoros Athanaileas, Theodora Varvarigou, “High Performance Computing in Biomedicine” in “Handbook of Research on Biocomputation and Biomedical Informatics: Case Studies and Applications”, IGI Global Publishers, 2009.
- [27] Dimosthenis Kyriazis, Konstantinos Tserpes, George Kousiouris, Andreas Menychtas, Gregory Katsaros, Theodora Varvarigou, “Data Aggregation and

- Analysis: A Grid-based approach for Medicine and Biology”, 4th Workshop on High Performance and Grid Computing in Medicine and Biology, The 2008 IEEE International Symposium on Parallel and Distributed Processing and Applications (ISPA), Sydney, Australia, 2008.
- [28] Dimosthenis Kyriazis, Peer Hasselmeyer, Konstantinos Tserpes, Andreas Menychtas and Theodora Varvarigou, “QoS-based Decision Services in Grids”, 12th Panhellenic Conference on Informatics (PCI), Samos, Greece, 2008.
- [29] Dionysiou D., Stamatakos G., Uzunoglu N. et al., “A four-dimensional simulation model of tumour response to radiotherapy in vivo: parametric validation considering radiosensitivity, genetic profile and fractionation”, Journal of Theoretical Biology, Volume 230, Issue 1, pp. 1-20, 2004.
- [30] Dionysiou D., Stamatakos G., Uzunoglu N., et al., “A computer simulation of in vivo tumour growth and response to radiotherapy: new algorithms and parametric results”, Computers in Biology and Medicine Journal, Volume 36, Issue 5, pp. 448-464, 2006.
- [31] E. Christensen et al., “Web Services Description Language (WSDL) 1.1,” W3C Note, 15 Mar. 2001, <http://www.w3.org/TR/wsdl>
- [32] EGEE Project: Enabling Grids for E-science, <http://www.eu-egee.org>
- [33] Elias Kourpas, “Grid Computing: Past, Present and Future”, IBM, <http://www.ibm.com/grid/pdf/innovperspective.pdf>
- [34] Erwin Laure, E. Fisher, S.M Fisher, Akos Frohner, C. Grandi, and Peter Kunszt, “Programming the Grid with gLite”, Computational Methods in Science and Technology, 12(1):33–45, 2006.
- [35] F. Gagliardi et al., “Building an infrastructure for scientific Grid computing: Status and goals of the EGEE project,” Philos. Trans.: Math. Phys. Eng. Sci., vol. 363, pp. 1729–1742, 2005.
- [36] Fladenmuller, A. et al., “A Hybrid QoS Management Scheme for Distributed Multimedia Applications”, in: Proceedings of Second Workshop on Protocols for Multimedia Systems PROMS'95, Salzburg Austria, October 1995, pp. 427-440.



- [37] Foster and C. Kesselman, “The Grid: Blueprint for a Future Computing Infrastructure”, Morgan Kaufmann Publishers, USA, 1999.
- [38] Foster I., Kesselman C., Tuecke S., “The Anatomy of the Grid: Enabling Scalable Virtual Organizations”, International Journal of Supercomputer Applications, Volume 15, Issue 3, pp. 200-222, 2001.
- [39] G. Wang, A. Chen, C. Wang, C. Fung, and S. Uczekaj, “Integrated Quality of Service (QoS) Management in Service-Oriented Enterprise Architectures”, In Proceedings of the 8th International IEEE Enterprise Distributed Object Computing Conference (EDOC), Monterey, California, September 2004.
- [40] Gartner Research, Understanding Hype Cycle, <http://www.gartner.com/pages/story.php.id.8795.s.8.jsp>
- [41] Geuer-Pollmann C, Claessens J. Web Services and Web Service Security Standards. Information Security Technical Report, Elsevier, 2005, 10(1)15:24.
- [42] gLite: Lightweight Middleware for Grid Computing, <http://glite.web.cern.ch/glite>
- [43] GRIA: Service Oriented Collaborations for Industry and Commerce, <http://www.gria.org>
- [44] GRID-APP: Ανάπτυξη Πιλοτικών Εφαρμογών Grid για την αξιοποίηση των Εθνικών Υποδομών Πλέγματος, [http://www.gsrt.gr/default.asp?V\\_ITEM\\_ID=4318](http://www.gsrt.gr/default.asp?V_ITEM_ID=4318)
- [45] Hellas Grid, <http://www.hellasgrid.gr>
- [46] Henry Chesborough and Richard Rosenbloom, the role of the business model in capturing value from innovation: evidence from Xerox Corporation’s technology spin-off companies, Industrial and Corporate Change, Volume 11, Number 3 pp529-555.
- [47] Hong-Linh Truong; Samborski, R.; Fahringer, T., "Towards a Framework for Monitoring and Analyzing QoS Metrics of Grid Services," e-Science and Grid Computing, 2006. e-Science '06. Second IEEE International Conference Dec. 2006.
- [48] [http://en.wikipedia.org/wiki/Multitier\\_architecture](http://en.wikipedia.org/wiki/Multitier_architecture)

- [49] Huedo E., Montero R., Llorente I., “Experiences on Adaptive Grid Scheduling of Parameter Sweep Applications”, Proceedings of the 12th Euromicro Conference on Parallel, Distributed and Network-based Processing (PDP’04). A Coruña, Spain, pp. 28-33, 2004
- [50] Introduction to the NextGRID Vision and Achievements, [http://www.nextgrid.org/download/publications/NextGRID\\_Architecture\\_White\\_Paper.pdf](http://www.nextgrid.org/download/publications/NextGRID_Architecture_White_Paper.pdf)
- [51] IRMOS - Interactive Realtime Multimedia Applications on Service Oriented Infrastructures, <http://www.irmosproject.eu>
- [52] IRMOS Project: ISONI Whitepaper, [http://www.irmosproject.eu/Files/IRMOS\\_WP6\\_7\\_ISONI\\_White\\_Paper\\_ALUD\\_USTUTT\\_v1\\_0.pdf](http://www.irmosproject.eu/Files/IRMOS_WP6_7_ISONI_White_Paper_ALUD_USTUTT_v1_0.pdf)
- [53] J. Pu, X. Zhang, and Z. Wu, “The Research on QoS for Grid Computing,” Proc. 2003 Int'l Conf. Comm. Techniques (ICCT '03), pp. 1711-1714, 2003.
- [54] K. Czajkowski, D. Ferguson, I. Foster, J. Frey, S. Graham, I. Sedukhin, D. Snelling, S. Tuecke, W. Vambenepe, Web Services Resource Framework (WSRF), Globus Alliance and IBM, 2005.
- [55] Kacsuk P. and Sipos G., “Multi-Grid, Multi-User Workflows in the P-GRADE Portal”, Journal of Grid Computing, Volume 3, Issues 3-4, pp. 221-238, 2005.
- [56] Keith Adams , Ole Agesen, A comparison of software and hardware techniques for x86 virtualization, Proceedings of the 12th international conference on Architectural support for programming languages and operating systems, October 21-25, 2006, San Jose, California, USA.
- [57] Konstantinos Dolkas, Dimosthenis Kyriazis, Andreas Menychtas, Theodora Varvarigou, “e-Business Applications on the Grid: A Toolkit for Centralized Workload Prediction and Access” published in Concurrency and Computation: Practice and Experience, Willey Interscience, 2006.
- [58] Konstantinos Tserpes, Dimosthenis Kyriazis, Andreas Menychtas and Theodora Varvarigou, “A Novel Mechanism for Provisioning of High-Level Quality of Service Information in Grid Environments”, to appear in Special Issue on

- Performance Evaluation of QoS-aware Heterogeneous Systems, European Journal of Operational Research, Elsevier, 2007.
- [59] Konstantinos Tserpes, Dimosthenis Kyriazis, Andreas Menychtas and Theodora Varvarigou, Fabrizio Silvestri, Domenico Laforenza, “An Open Architecture for QoS Information in Business Grids”, CoreGRID Symposium, August 27-28, 2007, Rennes, France.
- [60] L. Abeni and G. Buttazzo. Resource reservation in dynamic real-time systems. Journal of Real-Time Systems, 27(2):123–167, 1998.
- [61] L. Abeni, T. Cucinotta, G. Lipari, L. Marzario, and L. Palopoli. Qos management through adaptive reservations. Journal of Real-Time Systems, 29(2-3):131–155, 2005.
- [62] M. Surridge, S. Taylor, D. De Roure, and E. Zaluska, “Experiences with GRIA-Industrial Applications on a Web Services Grid”, in Proceedings of the First International Conference on e-Science and Grid Computing, IEEE Press, 2005.
- [63] Magretta, J., 2002. Why business models matter, Harvard Business Review (May), 86–92.
- [64] Mitchell, B., Mckee, P.: SLAs A Key Commercial Tool. In: Cunningham, P., Cunningham, M. (eds.) Innovation and the Knowledge Economy: Issues, Applications, Case Studies, IOS Press, Amsterdam (2006).
- [65] Newcomer, Eric; Lomow, Greg (2005). Understanding SOA with Web Services. Addison Wesley. ISBN 0-321-18086-0.
- [66] NextGRID Project, Integration Diagram, [http://grid.ece.ntua.gr/NextGRIDWiki/index.php/Integration\\_Diagram](http://grid.ece.ntua.gr/NextGRIDWiki/index.php/Integration_Diagram)
- [67] NextGRID Project, NextGRID Cookbook, [http://grid.ece.ntua.gr/NextGRIDWiki/index.php/NextGRID\\_Cookbook](http://grid.ece.ntua.gr/NextGRIDWiki/index.php/NextGRID_Cookbook)
- [68] NextGRID Project: Architecture for Next Generation Grids, <http://www.nextgrid.org>
- [69] Novotny J., Tuecke S., Welch V., "An Online Credential Repository for the Grid: MyProxy," Proceedings of the 10th International Symposium on High Performance Distributed Computing, New York, pp. 104-111, 2001.

- [70] OASIS Service Oriented Architecture Reference Model: [http://www.oasis-open.org/committees/tc\\_home.php?wg\\_abbrev=soa-rm](http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=soa-rm)
- [71] Olle Mulmo et al., EGEE Deliverable DJRA3.1, 2004-09-17, available online at: <https://edms.cern.ch/document/487004>
- [72] Open Grid Forum, [www.ogf.org](http://www.ogf.org)
- [73] Open Grid Services Architecture (OGSA), [www.ggf.org/documents/GFD.30.pdf](http://www.ggf.org/documents/GFD.30.pdf)
- [74] P. Bahl and V. N. Padmanabhan. Radar: An in-building user location and tracking system. In Proceedings of the IEEE Infocom 2000, volume 2, pages 775–84, Tel-Aviv, Israel, March 2000.
- [75] Palopoli, Luigi and Abeni, Luca and Cucinotta, Tommaso and Marzario, Luca and Lipari, Giuseppe (2007) Adaptive management of QoS in Open Systems. Technical Report DIT-07-003, Informatica e Telecomunicazioni, University of Trento.
- [76] Paul Timmers, Business Models for Electronic Markets, European Commission, DG III April 1998.
- [77] Philipp Masche et al., The increasing role of SLAs in B2B, Proceedings of the 2nd international conference on web information systems and technologies, April 2006, Setubal, Portugal.
- [78] Rajkumar Buyya, Chee Shin Yeo, and Srikumar Venugopal. Market-oriented cloud computing: Vision, hype, and reality for delivering it services as computing utilities. CoRR, (abs/0808.3558), 2008.
- [79] Ribba B, Colin T, Schnell S. 2006. A multiscale mathematical model of cancer, and its use in analyzing irradiation therapies. Theor Biol Med Mod 3:7.
- [80] S. Tuecke, K. Czajkowski, I. Foster, J. Frey, S. Graham, C. Kesselman, T. Maquire, T.Sandholm, D.Snelling, and P.Vanderbilt. Open Grid Service Infrastructure (OGSI). Technical report, OGSI-WG, Global Grid Forum, 2003. Version 1.0.
- [81] Service Level Agreement Zone, <http://www.sla-zone.co.uk>
- [82] Service-Oriented Infrastructure, SOA Working Group, <http://www.opengroup.org/projects/soa-soi>

- [83] SOAP Version 1.2 Part 1: Messaging Framework (Second Edition), W3C Recommendation 27 April 2007, <http://www.w3.org/TR/soap12-part1>
- [84] Spyridon Gogouvitis, George Kousiouris, Kleopatra Konstanteli, Theodoros Polychniatis, Andreas Menychtas, Dimosthenis Kyriazis, Theodora Varvarigou, “Realtime-enabled Workflow Management in Service Oriented Infrastructures”, 1st ACM International Workshop in Analysis and Retrieval of Events/Actions and Workflows in Video Streams (AREA), ACM Multimedia 2008, Vancouver, Canada, 2008.
- [85] Stamatakos G., “In silico oncology: a paradigm for clinically oriented living matter engineering”, Proceedings of the 3rd International Advanced Research Workshop on In Silico Oncology: Advances and Challenges, Istanbul, Turkey, pp. 7-9, 2008.
- [86] Stamatakos G., Dionysiou D., Zacharaki E. et al., “In silico radiation oncology: combining novel simulation algorithms with current visualization techniques”, Proceedings of the IEEE, Special Issue on “Bioinformatics: Advances and Challenges”, Volume 90, Issue 11, pp. 1764- 1777, 2002.
- [87] Subramani V., Ketimuthu R., Srinivasan S., et al., “Distributed job scheduling on computational grids using multiple simultaneous requests”, Proceedings of the 11th IEEE International Symposium on High Performance Distributed Computing, Edinburgh, Scotland, pp. 359-367, 2002.
- [88] Sudholta W., Baldridgea K., Abramson D., et al., "Application of grid computing to parameter sweeps and optimizations in molecular modeling", Future Generation Computer Systems, Volume 21, Issue 1, pp. 27-35, 2005.
- [89] The “ACGT: Advancing Clinicogenomic Trials on Cancer” Project (FP6-2005-IST-026996), URL: <http://eu-acgt.org/acgt-for-you/researchers/in-silico-oncology.html>
- [90] The “ContraCancrum: Clinically Oriented Translational Cancer Multilevel Modelling” Project:
- [91] The GLUE schema, <http://glueschema.forge.cnaf.infn.it>
- [92] The grid: blueprint for a new computing infrastructure. By Ian Foster, Carl Kesselman Published 2003 ISBN:1558609334.

- [93] The South-Eastern Europe Virtual Organization (SEE-VO), <http://www.egee-see.org/see-vo.php?language=en>
- [94] Theodoros Athanaileas, Andreas Menychtas, Dimitra Dionysiou, Dimosthenis Kyriazis, Dimitra Kaklamani, Theodora Varvarigou, Nikolaos Uzunoglu and Georgios Stamatakos, “Exploiting Grid Technologies for the Simulation of Clinical Trials In Silico: the Paradigm of In Silico Radiation Oncology”, Techniques and Architectures for Computer Simulation on SIMULATION: Transactions of the Society for Modeling and Simulation International, 2008.
- [95] Theodoros Athanaileas, Andreas Menychtas, Dimitra Dionysiou, Georgios Stamatakos, Dimitra Kaklamani, Theodora Varvarigou, Nikolaos Uzunoglu, “A Grid-Enabled Toolkit for In Silico Oncology Simulations”, SIMUTools 2008, Marseille, France.
- [96] Thomas Erl, “SOA Principles of Service Design (The Prentice Hall Service-Oriented Computing Series), Chapter 5: Understanding Design Principles”, Prentice Hall/PearsonPTR, ISBN: 0132344823.
- [97] Titz B and Jeraj R 2008. An imaging-based tumour growth and treatment response model: investigating the effect of tumour oxygenation on radiation therapy response. Phys Med Biol 53: 4471-4488.
- [98] Urgaonkar B., Pacifici G., Shenoy P., et al., “An analytical model for multi-tier internet services and its applications”, Proceedings of the 2005 ACM SIGMETRICS international conference on Measurement and modeling of computer systems, Banff, Alberta, Canada, pp. 291-302, 2005.
- [99] Vler Sander, William A. Adamson, Ian Foster, “End-to-End Provision of Policy Information for Network QoS”, Proceedings of the Tenth IEEE Symposium on HPDC, August 2001.
- [100] W. Leinberger, V. Kumar, “Information Power Grid: The new frontier in parallel computing?” IEEE Concur., Vol. 7, No. 4, pp. 75-84, October-December 1999.
- [101] Web Services Agreement Specification (WS-Agreement), OGF, <http://www.gridforum.org/Meetings/GGF11/Documents/draft-ggf-graap-agreement.pdf>

- [102] Weng C., Lu X., “Heuristic scheduling for bag-of-tasks applications in combination with QoS in computational grid”, *Future Generation Computer Systems*, Volume 21, Issue 2, pp. 271-280, 2005.
- [103] Werner-Wasik M., Scott C. , Nelson D., “Final report of a phase I/II trial of hyperfractionated and accelerated hyperfractionated radiation therapy with carmustine for adults with supratentorial malignant gliomas: Radiation therapy oncology group study 83-02”, *Cancer Journal*, Volume 77, Issue 8, pp. 1535 – 1543, 1998.
- [104] Xia Y., Wang H., Xu C., et al., “Stochastic modelling and quality evaluation of workflow systems based on QWF-nets”, *3rd International Conference on Computational Science (ICCS)*, pp. 988-995, 2006.