



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ

ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΕΠΙΚΟΙΝΩΝΙΩΝ, ΗΛΕΚΤΡΟΝΙΚΗΣ & ΣΥΣΤΗΜΑΤΩΝ
ΠΛΗΡΟΦΟΡΙΚΗΣ

Τεχνολογίες Πλέγματος με έμφαση στις δυναμικές Ροές Εργασίας και στην Ποιότητα Υπηρεσιών

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Δημοσθένης Π. Κυριαζής

Αθήνα, Ιούλιος 2007



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΕΠΙΚΟΙΝΩΝΙΩΝ, ΗΛΕΚΤΡΟΝΙΚΗΣ & ΣΥΣΤΗΜΑΤΩΝ
ΠΛΗΡΟΦΟΡΙΚΗΣ

Τεχνολογίες Πλέγματος με έμφαση στις δυναμικές Ροές Εργασίας και στην Ποιότητα Υπηρεσιών

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Δημοσθένης Π. Κυριαζής

Συμβουλευτική Επιτροπή : Θεοδώρα Α. Βαρβαρίγου

Εμμανουήλ Ν. Πρωτονοτάριος

Μιχαήλ Ε. Θεολόγου

Εγκρίθηκε από την επταμελή εξεταστική επιτροπή την 27^η Ιουλίου 2007.

.....
Θ.Α. Βαρβαρίγου
Αναπληρώτρια Καθηγήτρια
Ε.Μ.Π.

.....
Ε.Ν. Πρωτονοτάριος
Καθηγητής Ε.Μ.Π.

.....
Μ.Ε. Θεολόγου
Καθηγητής Ε.Μ.Π.

.....
Ν.Κ. Ουζούνoglou
Καθηγητής Ε.Μ.Π.

.....
Γ.Ι. Στασινόπουλος
Καθηγητής Ε.Μ.Π.

.....
Α. Γ. Σταφυλοπάτης
Καθηγητής Ε.Μ.Π.

.....
Ε. Βαρβαρίγος
Καθηγητής Πανεπιστημίου
Πατρών

Αθήνα, Ιούλιος 2007

.....
Δημοσθένης Π. Κυριαζής

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

Copyright © Δημοσθένης Π. Κυριαζής, 2007.

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.

ΠΡΟΛΟΓΟΣ

Η διδακτορική διατριβή που παρουσιάζεται στις επόμενες σελίδες εκπονήθηκε από το Νοέμβριο του 2003 μέχρι τον Ιούλιο του 2007, στο εργαστήριο Τηλεπικοινωνιών του τομέα Επικοινωνιών, Ηλεκτρονικής και Συστημάτων Πληροφορικής, στη Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του Εθνικού Μετσόβιου Πολυτεχνείου.

Κατά την διάρκεια της εκπόνησης αυτής της διατριβής, είχα την ευκαιρία να ασχοληθώ με αρκετά ενδιαφέροντα επιστημονικά θέματα που αφορούν κυρίως στους τομείς της προδιαγραφής, του σχεδιασμού, της υλοποίησης και του ελέγχου υποδομών πλέγματος 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 Μιχαήλ Θεολόγου.

Επίσης, θα ήθελα να ευχαριστήσω όλους τους συναδέλφους μου στην ερευνητική ομάδα με τους οποίους συνεργάστηκα άψογα και επιτυχώς όλα αυτά τα χρόνια.. Ιδιαίτερες ευχαριστίες ωστόσο θα ήθελα να απευθύνω στους στενούς μου συνεργάτες και κυρίως στους Κωνσταντίνο Χριστογιάννη, Ανδρέα Μενύχτα, Κωνσταντίνο Τσερπέ, Δημήτριο Χάλκο, Αντώνιο Λίτκε, Νικόλαο και Αναστάσιο Δουλάμη με τους οποίους μοιραστήκαμε τις πάρα πολλές ώρες της ερευνητικής εργασίας .

Τέλος, θα ήθελα να ευχαριστήσω τους γονείς μου και την αδερφή μου που πίστεψαν σε εμένα και τις επιλογές μου.

Δημοσθένης Π. Κυριαζής

Ιούλιος 2007

Πίνακας Περιεχομένων

Περίληψη	xv
Abstract	xvii
1 Εισαγωγή	1
1.1 Ορισμός του Πλέγματος	1
1.2 Όραμα και Ανοιχτά Τεχνολογικά Θέματα	3
1.3 Συνεισφορά – Καινοτομία Διατριβής	7
1.4 Οργάνωση της Διατριβής	10
2 Γενικά για το περιβάλλον Πλέγματος	13
2.1 Ιστορικά στοιχεία - από το Διαδίκτυο στο Πλέγμα	13
2.2 Χαρακτηριστικά	17
2.3 Οι Γενιές των Συστημάτων Πλέγματος	19
2.4 Όροι Πλέγματος	21
2.4.1 <i>To Meta-computing</i>	21
2.4.2 <i>Μη Υπολογιστικοί Πόροι</i>	22
2.4.3 <i>Εικονοποίηση</i>	23
2.4.4 <i>Υπηρεσίες Δικτύου</i>	24
2.4.5 <i>Εικονικοί Οργανισμοί</i>	25
2.4.6 <i>Συμβόλαια Παροχής Υπηρεσιών</i>	27
2.5 Ταξινόμηση Υποδομών Πλέγματος	33
2.6 Σύγκριση Τεχνολογίας Πλέγματος με άλλες Τεχνολογίες	35
2.7 Πλεονεκτήματα	37
2.8 Τομείς που ευνοούνται από την τεχνολογία Πλέγματος	39
3 Δυναμικές Ροές Εργασίας	41
3.1 Ορισμοί	42
3.2 Κατηγορίες	43
3.3 Καθορισμός Συγκεκριμένης Ροής Εργασίας	50
3.4 Γλώσσα Περιγραφής	50

3.5	Η γλώσσα OWL-WS	52
3.5.1	Χρήσιμοι ορισμοί	52
3.5.2	Τα βασικά στοιχεία της γλώσσας OWL-S	53
3.5.3	Το μοντέλο της γλώσσας OWL-WS	55
3.5.4	Οι παράμετροι των προφίλ	61
4	Αρχιτεκτονική Σχεδίαση	67
4.1	Παράμετροι Σχεδίασης	67
4.2	Υπηρεσιοστρεφής Αρχιτεκτονική	77
4.2.1	Γενικά για τις υπηρεσίες	77
4.2.2	Διαδικτυακές Υπηρεσίες	78
4.2.3	Πρότυπο Υπηρεσιοστρεφούς Αρχιτεκτονικής	79
4.2.4	Γλώσσα Περιγραφής Διαδικτυακών υπηρεσιών	81
4.2.5	Πρωτόκολλο Επικοινωνίας	81
4.3	Αρχιτεκτονική Ανοιχτών Υπηρεσιών Πλέγματος	83
4.3.1	Στόχοι της Αρχιτεκτονικής	83
4.3.2	Περιγραφή της Αρχιτεκτονικής	85
4.4	Καινοτόμα Αρχιτεκτονική Πλέγματος Επόμενης Γενιάς	90
4.4.1	Μοντέλο Εικονικής Υποδομής	91
4.4.2	Συγκεκριμένα Στοιχεία Αρχιτεκτονικής	97
4.4.3	Αρχιτεκτονική Σχεδίαση του Μοντέλου Εικονικής Υποδομής	125
5	Ποιότητα Υπηρεσίας	129
5.1	Γενικά Στοιχεία	129
5.1.1	Η Ποιότητα Υπηρεσίας σε περιβάλλον Πλέγματος	130
5.1.2	Από το Ακαδημαϊκό στο Επιχειρηματικό Πλέγμα	132
5.1.3	Προσέγγιση Λειτουργίας του Επιχειρηματικού Μοντέλου του Πλέγματος	133
5.2	Παρούσα κατάσταση - Σχετικές εργασίες	139
5.3	Καινοτόμος Κατηγοριοποίηση Παραμέτρων Ποιότητας Υπηρεσίας	141
5.4	Κατηγορίες απαιτήσεων	143
5.4.1	Απαιτήσεις για το Κόστος	144
5.4.2	Απαιτήσεις για τη Διαθεσιμότητα	145

5.4.3	<i>Απαιτήσεις για το Χρόνο Εκτέλεσης.....</i>	<i>146</i>
5.4.4	<i>Κατηγορίες Απαιτήσεων για όλες τις Παραμέτρους</i>	<i>147</i>
5.5	<i>Μοντέλα Δρομολόγησης Εργασιών και Εκτίμησης Ποιότητας Υπηρεσίας</i>	<i>148</i>
5.5.1	<i>Αρχιτεκτονική Υποδομής.....</i>	<i>148</i>
5.5.2	<i>Παράμετροι που απαιτούνται για τη Δρομολόγηση</i>	<i>151</i>
5.5.3	<i>Δρομολόγηση και παρεχόμενη Ποιότητα Υπηρεσίας</i>	<i>154</i>
5.5.4	<i>Διαχείριση Πόρων και Δρομολόγηση.....</i>	<i>155</i>
5.5.5	<i>Δίκαιη Δρομολόγηση.....</i>	<i>155</i>
6	<i>Απεικόνιση Δυναμικών Ροών Εργασίας και Επιλογή Υπηρεσιών</i>	<i>161</i>
6.1	<i>Παρούσα Κατάσταση – Σχετικές Εργασίες.....</i>	<i>161</i>
6.2	<i>Καινοτόμος Μηχανισμός Απεικόνισης Δυναμικών Ροών Εργασίας και Επιλογής Υπηρεσιών Πλέγματος.....</i>	<i>164</i>
6.3	<i>Προσδιορισμός του Προβλήματος και Λύση</i>	<i>168</i>
6.3.1	<i>Αντικατάσταση Διαθεσιμότητας, Χρόνου και Κόστους από ένα Κοινό Μέγεθος</i>	<i>168</i>
6.3.2	<i>Ανάγκη Τροποποίησης του Λόγου Κόστους και Χρόνου / Διαθεσιμότητα</i>	<i>169</i>
6.3.3	<i>Τροποποιημένος Λόγος Κόστους και Χρόνου / Διαθεσιμότητα</i>	<i>170</i>
6.3.4	<i>Η Συνάρτηση Μετατροπής.....</i>	<i>171</i>
6.4	<i>Καινοτόμος Αλγόριθμος Απεικόνισης Δυναμικών Ροών Εργασίας και Επιλογής Υπηρεσιών για Διαχείριση της Ποιότητας Υπηρεσίας</i>	<i>177</i>
6.5	<i>Καινοτόμο Σχήμα Καθορισμού Περισσότερων Συγκεκριμένων Ροών Εργασίας.....</i>	<i>185</i>
6.5.1	<i>Αρχιτεκτονική</i>	<i>186</i>
6.5.2	<i>Λειτουργικότητα.....</i>	<i>190</i>
6.6	<i>Ενσωμάτωση του Μηχανισμού Απεικόνισης στο μοντέλο GridVIM.....</i>	<i>193</i>
6.7	<i>Σύγκριση Διαμεσολαβητή και Μηχανισμού Απεικόνισης.....</i>	<i>197</i>
7	<i>Αξιολόγηση του Προτεινόμενου Μηχανισμού</i>	<i>201</i>
7.1	<i>Εφαρμογή</i>	<i>201</i>
7.2	<i>Σύστημα Δοκιμών</i>	<i>202</i>
7.3	<i>Αξιολόγηση</i>	<i>205</i>
7.3.1	<i>1^η Προσομοίωση.....</i>	<i>207</i>
7.3.2	<i>2^η Προσομοίωση.....</i>	<i>215</i>

7.3.3	3 ^η Προσομοίωση.....	221
8	Σύνοψη Κυριότερων Συνεισφορών - Ανοιχτά Ερευνητικά Θέματα.....	237
8.1	Σύνοψη Διατριβής και Κυριότερων Συνεισφορών.....	237
8.2	Ανοιχτά Ερευνητικά Θέματα.....	239
8.2.1	Στρατηγικές Εξαρτήσεις στις Ροές Εργασίας	239
8.2.2	Μηχανισμός Καθορισμού Βέλτιστου Τρόπου Εκτέλεσης Εργασιών	240
8.2.3	Ο Εικονικός Υπολογιστής ως Υπηρεσία	241
	Γλωσσάριο	243
	Βιβλιογραφικές Αναφορές.....	245

Σχήματα

Σχήμα 1: Ερευνητικές Προτεραιότητες	5
Σχήμα 2: Αρχιτεκτονική Globus	17
Σχήμα 3: Σχέσεις όρων σε συμβόλαια παροχής υπηρεσιών	31
Σχήμα 4: Ανάλυση ροών εργασίας από την ροή της εφαρμογής.....	45
Σχήμα 5: Ορισμοί ροής εργασίας.....	47
Σχήμα 6: Βασικά Στοιχεία της OWL-S	55
Σχήμα 7: Αναπαράσταση συγκεκριμένης ροής εργασίας στις OWL-S, OWL-WS	57
Σχήμα 8: Αναπαράσταση αφηρημένης ροής εργασίας στις OWL-S, OWL-WS.....	59
Σχήμα 9: Προφίλ OWL-S	61
Σχήμα 10: Ανταλλαγή ρόλων μεταξύ διαδικτυακών υπηρεσιών κατά τη διάρκεια επικοινωνίας	79
Σχήμα 11: Δομή Υπηρεσιοστρεφούς αρχιτεκτονικής.....	79
Σχήμα 12: Χρήση SOAP πρωτοκόλλου σε εφαρμογές.....	83
Σχήμα 13: Αρχιτεκτονική OGSA	85
Σχήμα 14: Αρχιτεκτονική του μοντέλου GridVIM	93
Σχήμα 15: Μοντέλο GridVIM.....	93
Σχήμα 16: Καθορισμός προτεραιοτήτων υπηρεσιών	95
Σχήμα 17: Εύρεση υπηρεσιών	95
Σχήμα 18: Επιλογή υπηρεσιών	95
Σχήμα 19: Δομικά στοιχεία του GridVIM.....	97
Σχήμα 20: Περιπτώσεις χρήσης.....	99
Σχήμα 21: Βασικά στοιχεία και λειτουργίες δομής GridVIM.....	101
Σχήμα 22: Ερωτήσεις προφίλ	105
Σχήμα 23: Δυνατότητες προφίλ.....	105
Σχήμα 24: Αντιστοίχιση προφίλ στις ερωτήσεις	107
Σχήμα 25: Αποτελέσματα αντιστοίχισης	109
Σχήμα 26: Υπηρεσία μητρώου	111
Σχήμα 27: Περιγραφή μοντέλου	113
Σχήμα 28: Μοντέλο διαμεσολάβησης	117
Σχήμα 29: Ανταλλαγή δεδομένων μεταξύ καταναλωτή και διαμεσολαβητή	117

Σχήμα 30: Εσωτερική αρχιτεκτονική διαμεσολαβητή	119
Σχήμα 31: Μοντέλο δεδομένων	121
Σχήμα 32: Αρχιτεκτονική του μοντέλου εικονικής υποδομής και αλληλεπιδράσεις	125
Σχήμα 33: Διάγραμμα αλληλεπιδράσεων	125
Σχήμα 34: Σχηματική απεικόνιση υλοποίησης διεπαφών και αλληλεπιδράσεων	127
Σχήμα 35: Από το ακαδημαϊκό στο επιχειρηματικό πλέγμα	133
Σχήμα 36: Ταξινόμηση παραμέτρων ποιότητας υπηρεσίας	141
Σχήμα 37: Αρχιτεκτονική περιγραφή υποδομής	149
Σχήμα 38: Συνοπτική παρουσίαση μηχανισμού απεικόνισης ροών εργασίας	165
Σχήμα 39: Γραφική αναπαράσταση μεταβολής κόστους	173
Σχήμα 40: Γραφική αναπαράσταση μεταβολής κόστους για συντελεστή βάρους ίσο με 0.8	173
Σχήμα 41: Γραφική αναπαράσταση μεταβολής κόστους για συντελεστή βάρους ίσο με 0.2	173
Σχήμα 42: Αρχιτεκτονική συστήματος	187
Σχήμα 43: Ροή λογικής μεταξύ των στοιχείων	193
Σχήμα 44: Διάγραμμα κλάσεων του μηχανισμού απεικόνισης ροών εργασίας	197
Σχήμα 45: Σενάριο παραγωγής τρισδιάστατης εικόνας	203
Σχήμα 46: Αφηρημένη ροή εργασίας και αρχικοί τύποι & υποστάσεις υπηρεσιών	205
Σχήμα 47: Συγκεκριμένη ροή εργασίας για την 1 ^η προσομοίωση	211
Σχήμα 48: Μεταβολή τιμών διαθεσιμότητας, κόστους και χρόνου με βάση τις επιλογές που πραγματοποιούνται κατά τις επαναλήψεις του αλγορίθμου για την 1 ^η προσομοίωση	213
Σχήμα 49: Συγκεκριμένη ροή εργασίας για τη 2 ^η προσομοίωση	219
Σχήμα 50: Μεταβολή τιμών διαθεσιμότητας, κόστους και χρόνου με βάση τις επιλογές που πραγματοποιούνται κατά τις επαναλήψεις του αλγορίθμου για τη 2 ^η προσομοίωση	219
Σχήμα 51: Αρχική έκδοση συγκεκριμένης ροής εργασίας	225
Σχήμα 52: 1 ^η Εναλλακτική έκδοση συγκεκριμένης ροής εργασίας	229
Σχήμα 53: 2 ^η Εναλλακτική έκδοση συγκεκριμένης ροής εργασίας	233
Σχήμα 54: Μεταβολή τιμών χαρακτηρισμού, κόστους και χρόνου με βάση τις επιλογές που πραγματοποιούνται κατά τις επαναλήψεις του αλγορίθμου για τις τρεις εκδόσεις συγκεκριμένων ροών εργασίας	233

Πίνακες

Πίνακας 1: Προδιαγραφές των υπολογιστών της φάρμας (cluster).....	202
Πίνακας 2: Τιμές που ελήφθησαν αρχικά για τις υποστάσεις των υπηρεσιών	207
Πίνακας 3: Επαναλήψεις αλγορίθμου για την 1η προσομοίωση	209
Πίνακας 4: Επαναλήψεις αλγορίθμου στο μηχανισμό χρονοπρογραμματισμού για την 1 ^η προσομοίωση	210
Πίνακας 5: Επαναλήψεις αλγορίθμου για τη 2 ^η προσομοίωση.....	216
Πίνακας 6: Επαναλήψεις αλγορίθμου στο μηχανισμό χρονοπρογραμματισμού για τη 2 ^η προσομοίωση	217
Πίνακας 7: Τιμές που ελήφθησαν αρχικά για τις υποστάσεις των υπηρεσιών	222
Πίνακας 8: Επαναλήψεις αλγορίθμου για την αρχική έκδοση συγκεκριμένης ροής εργασίας.....	224
Πίνακας 9: Επαναλήψεις αλγορίθμου στο μηχανισμό χρονοπρογραμματισμού για την αρχική συγκεκριμένη ροή εργασίας.....	225
Πίνακας 10: Επαναλήψεις Αλγορίθμου για την 1 ^η εναλλακτική έκδοση συγκεκριμένης ροής εργασίας	228
Πίνακας 11: Επαναλήψεις αλγορίθμου στο μηχανισμό χρονοπρογραμματισμού για την 1 ^η εναλλακτική έκδοση συγκεκριμένης ροής εργασίας.....	228
Πίνακας 12: Επαναλήψεις αλγορίθμου για την 2 ^η εναλλακτική έκδοση συγκεκριμένης ροής εργασίας	231
Πίνακας 13: Επαναλήψεις αλγορίθμου στο μηχανισμό χρονοπρογραμματισμού για τη 2 ^η εναλλακτική ροή εργασίας	231
Πίνακας 14: Σύνοψη επιπέδων προσφερόμενης ποιότητας υπηρεσίας για την αρχική και τις εναλλακτικές ροές εργασίας.....	235

Περίληψη

Η ανάπτυξη των ετερογενών και κατανεμημένων περιβαλλόντων, όπως τα περιβάλλοντα πλέγματος, καθιστά εφικτή την επίλυση υπολογιστικά εντατικών προβλημάτων με αξιόπιστο και οικονομικό τρόπο. Στο πλαίσιο αυτό και καθώς η έννοια των δυναμικών Ροών Εργασίας αναφέρεται στην εκτέλεση εφαρμογών με αυξημένη πολυπλοκότητα (η οποία περιγράφεται στις αντίστοιχες ροές των εφαρμογών), κρίνεται αναγκαία η ανάλυση και διασφάλιση της Ποιότητας Υπηρεσιών που παρέχεται κατά τη φάση εκτέλεσης των ροών εργασίας σε περιβάλλον πλέγματος. Βάσει αυτού παρέχεται ένας πλήρης ορισμός και μια εμπεριστατωμένη κατηγοριοποίηση της έννοιας της Ροής Εργασίας και αναλύονται σχετικά θέματα όπως για παράδειγμα οι γλώσσες περιγραφής αυτών.

Στην παρούσα διατριβή παρουσιάζεται ένας μηχανισμός απεικόνισης των διαδικασιών που περιλαμβάνονται στις ροές εργασίας σε υπηρεσίες που παρέχονται μέσα από περιβάλλοντα πλέγματος. Η προαναφερθείσα απεικόνιση ολοκληρώνεται με τρόπο που να εξασφαλίζει ταυτόχρονα την προσφερόμενη ποιότητα υπηρεσιών για το σύνολο της ροής εργασίας με βάση τις παραμέτρους και τις προτιμήσεις που ορίζονται από τους χρήστες. Το τελευταίο θεωρείται προϋπόθεση για την υιοθέτηση διαφόρων επιχειρηματικών μοντέλων και εφαρμογών σε περιβάλλον πλέγματος καθώς με την απεικόνιση επιτυγχάνεται αξιολόγηση των εναλλακτικών στρατηγικών και παρακολούθηση των μεταβολών / παραβιάσεων που ενδεχομένως να συμβούν και που έχουν σημαντική επίδραση στις στρατηγικές, στις μεθοδολογίες και στη δομή των επιχειρηματικών διαδικασιών. Λόγω των παραπάνω, ο

μηχανισμός απεικόνισης δυναμικών ροών εργασίας αποτελεί βασικό μέσο διασφάλισης ποιότητας υπηρεσίας καθώς με τον τρόπο αυτό δύναται να εκτιμηθεί, να υπολογιστεί και να καθοριστεί η τελική επιλογή των διαθέσιμων τύπων και υποστάσεων υπηρεσιών προκειμένου να παρασχεθεί ποιότητα υπηρεσιών από ένα σύνολο παρόχων.

Επιπρόσθετα παρουσιάζεται και αναλύεται η Αρχιτεκτονική Πλέγματος Επόμενης Γενιάς με έμφαση σε συγκεκριμένα στοιχεία αυτής που επιτρέπουν το δυναμικό χαρακτήρα των ροών εργασίας. Βάσει αυτού παρουσιάζεται το Μοντέλο Εικονικής Υποδομής, ένα μοντέλο που στοχεύει στην αφαίρεση των επιχειρησιακών διεργασιών και των αλληλεπιδράσεων με την υποδομή πλέγματος κάνοντας χρήση προσαρμοσμένων ροών εργασίας και δυναμική δέσμευση υπηρεσιών, η οποία περιλαμβάνει σημασιολογική εύρεση και επιλογή αυτών.

Τα αποτελέσματα προσομοίωσης που παρατίθενται αξιολογούν και αποδεικνύουν τη λειτουργία και αποδοτικότητα του προτεινόμενου μηχανισμού λαμβάνοντας υπόψη τους περιορισμούς ποιότητας και τις προτιμήσεις του χρήστη. Το παραπάνω καθιστά εφικτή την ενσωμάτωση και χρήση του μηχανισμού σε οποιοδήποτε ετερογενές περιβάλλον και, ειδικότερα, σε υποδομές πλέγματος που στοχεύουν στην εισαγωγή της γνώσης αναφορικά με την ποιότητα των παρεχόμενων υπηρεσιών κατά τη διαδικασία επιλογής υπηρεσιών για τις διεργασίες των δυναμικών ροών εργασίας.

Abstract

The advent of heterogeneous and distributed environments, such as Grid environments, made feasible the solution of computational intensive problems in a reliable and cost-effective way. As workflow systems carry out more complex and mission-critical applications, Quality of Service (QoS) analysis serves to ensure that each application meets user requirements. In that frame and considering the QoS provision aspect as fundamental for enabling the adoption of different business models and Grid applications to become QoS compliant, a novel mechanism is presented. Based on this a complete definition and a thorough categorisation of the Workflows is presented and various relative subjects are analyzed such as the workflow description languages.

This thesis presents a Workflow Mapping Mechanism that allows the mapping of workflow processes to Grid provided services assuring at the same time end-to-end provision of QoS based on the user-defined parameters and preferences. The latter is considered as precondition for the adoption of different business models and objectives in Grid environments since the mechanism achieves the evaluation of the alternative strategies and monitoring of any changes / violations that may occur and which will have an important impact on the strategies, methodologies, and structure of business processes. Therefore, the workflow mapping mechanism is an integral part of the QoS provisioning, and especially the end-to-end Quality of Service, since this is the only way to estimate, calculate and conclude

to the mapping of workflows and the selection of the available service types and instances in order to deliver an overall quality of service across a federation of providers.

Furthermore, the Next Generation Grid Architecture is presented and analyzed with emphasis on concrete elements and components that enable workflows to be described and executed in a dynamic way. Based on this, a Grid Virtual Infrastructure Model is presented, which aims into abstracting the operational processes and interactions from the Grid infrastructure by making use of adaptive workflows and dynamic service allocation that includes service discovery and selection.

Consequently, the experimental results are provided that demonstrate and evaluate the performance and effectiveness of the proposed mechanism for a real Grid application scenario (3D image rendering), while the presented set of parameters (application, resource, user-defined) and their integration into the mechanism was revealed. Given the experimental results, it appears that the mechanism makes the selections and as a result defines the concrete workflow taking into account the user constraints and preferences. The experiments showed promising results and therefore the performance of the mechanism is considered to be well established allowing the adoption of it in any heterogeneous and especially Grid system that seeks to bring QoS knowledge within the workflow mapping process.

1

Εισαγωγή

Στο παρόν κεφάλαιο παρατίθεται αρχικά ο ορισμός του Πλέγματος («Grid») και στη συνέχεια αναλύεται η οργάνωση της παρούσης Διατριβής και περιγράφεται συνοπτικά το περιεχόμενο των κεφαλαίων.

1.1 Ορισμός του Πλέγματος

Η τεχνολογία των υπολογιστών και των τηλεπικοινωνιών γνωρίζει την κορύφωση 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 πλέγματος 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 είναι μια προσέγγιση της σύστασης δυναμικά δομημένων περιβαλλόντων, χρησιμοποιώντας υπολογιστικούς πόρους που είναι διεσπαρμένοι τόσο γεωγραφικά όσο και οργανωτικά. Ο όρος πλέγμα 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 πλέγματος έτσι όπως δίνεται από την IBM [1]:

Πλέγμα 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 πλέγμα είναι ένα είδος παράλληλου και κατανεμημένου συστήματος που δίνει τη δυνατότητα να διαμοιρασμού, να επιλογής και συγκέντρωσης πόρων που κατανέμονται σε πολλαπλές λειτουργικές περιοχές βάσει της διαθεσιμότητας των πόρων, της χωρητικότητας, της επίδοσης, του κόστους και τις απαιτήσεις ποιότητας υπηρεσίας που καθορίζονται από το χρήστη.

Από τα παραπάνω, γίνεται φανερό ότι ο όρος πλέγμα περιλαμβάνει το σύνολο της υποδομής, καθώς και των απαραίτητων υπηρεσιών για τη δημιουργία ενός ενιαίου υπερ-υπολογιστικού 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 περιβάλλον πλέγματος 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.

1.2 Όραμα και Ανοιχτά Τεχνολογικά Θέματα

Καθώς η έρευνα σε υπηρεσίες και υποδομές πλέγματος γνωρίζει την κορύφωση της τόσο σε Ευρωπαϊκό όσο και σε παγκόσμιο επίπεδο, έχουν διατυπωθεί διάφορες εργασίες από ομάδες εμπειρων επιστημόνων αναφορικά με το όραμα και τα ανοιχτά τεχνολογικά θέματα των υποδομών πλέγματος. Στην ενότητα αυτή παρατίθενται στοιχεία από ερευνητές που συγκροτούν μια ομάδα (Next Generation Grid Expert Group) που αναλαμβάνει να οριοθετήσει το όραμα και τα ανοιχτά θέματα των υποδομών πλέγματος επόμενης γενιάς [72]. Επιπρόσθετα αναφέρονται τα θέματα στα οποία απαντά η συγκεκριμένη διατριβή και προτείνοντας καινοτόμες λύσεις στο συγκεκριμένο χώρο.

Αναφορικά με το όραμα για το περιβάλλον πλέγματος, αυτό εκφράζεται από την προαναφερθείσα ομάδα ερευνητών ως μια μελλοντική υποδομή (που οριοθετείται στην

επόμενη δεκαετία) για την παροχή υπηρεσιών και λογισμικού και ονομάζεται «*Service Oriented Knowledge Utilities (SOKU)*». Η έννοια αυτή παρουσιάζει τα παρακάτω χαρακτηριστικά:

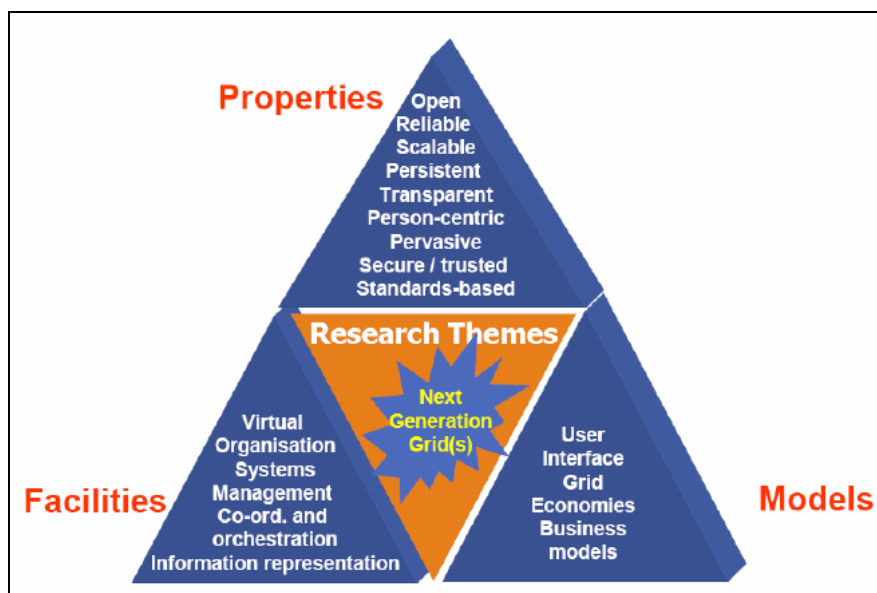
- Είναι ένας ευέλικτος, ισχυρός και οικονομικά αποδοτικός τρόπος, λειτουργίας και παροχής τεχνολογικών και επικοινωνιακών λύσεων προς χρήση από τις επιχειρήσεις, την επιστήμη και την κοινωνία
- Στηρίζεται στις υπάρχουσες επιχειρηματικές πρακτικές, τις τάσεις και τις νέες τεχνολογίες
- Παρέχει τους κανόνες και τις μεθόδους για το συνδυασμό των υπηρεσιών προς δημιουργία ενός οικοσυστήματος το οποίο αναμένεται να προωθήσει τις επιχειρηματικές συνεργασίες
- Παρουσιάζει τα οφέλη της μειωμένης πολυπλοκότητας, των χαμηλότερων εξόδων και της ευρύτερης διαθεσιμότητας των χρήσιμων υπηρεσιών

Τεχνολογικά, το όραμα SOKU στηρίζεται στη φυσική εξέλιξη και των συνδυασμό κυρίως των εννοιών των Διαδικτυακών Υπηρεσιών (Web Services), των τεχνολογιών Πλέγματος (Grid technologies) και του Σημασιολογικού Διαδικτύου (Semantic Web). Βάσει των παραπάνω, το όνομα αντικατοπτρίζει τις ακόλουθες βασικές έννοιες:

- *Service Oriented*: Η αρχιτεκτονική περιλαμβάνει τις υπηρεσίες που μπορούν αρχικοποιηθούν και να συγκεντρωθούν δυναμικά και ως εκ τούτου η δομή, η συμπεριφορά και η θέση του λογισμικού δύναται να αλλάζει κατά τη διάρκεια εκτέλεσης των εφαρμογών
- *Knowledge*: Οι υπηρεσίες SOKU εμπεριέχουν πρόσθετη πληροφορία («σημασιολογική») για την πραγματοποίηση αυτοματοποιημένων και προηγμένων λειτουργιών, το οποίο επιτρέπει την παροχή υπηρεσιών υψηλού επιπέδου στο χρήστη

- *Utility*: Πρόκειται για μια άμεσα χρησιμοποιήσιμη υπηρεσία με σαφώς καθιερωμένη και συγκεκριμένη λειτουργικότητα, απόδοση και αξιοπιστία, γεγονός που επεξηγεί την έμφαση που δίνεται στις ανάγκες των χρηστών.

Στο σχήμα που ακολουθεί (Σχήμα 1) παρουσιάζονται οι ερευνητικές προτεραιότητες όπως αυτές παρουσιάζονται στην αναφορά [72]. Επίσης οι προτεραιότητες αυτές κατατάσσονται σε τρεις κύριες κατηγορίες: Ιδιότητες (Properties), Υποδομές (Facilities) και Μοντέλα (Models).



Σχήμα 1: Ερευνητικές Προτεραιότητες

Βάσει του παραπάνω σχήματος η διατριβή προτείνει καινοτόμες λύσεις σε θέματα που αφορούν στη διαχείριση των συστημάτων και την ενορχήστρωση των υπηρεσιών πλέγματος με στόχο την παροχή ποιότητας υπηρεσιών. Το τελευταίο κρίνεται σημαντικό για την υιοθέτηση και εφαρμογή επιχειρηματικών μοντέλων σε περιβάλλοντα πλέγματος (το οποίο επίσης συναντάται στο παραπάνω σχήμα) καθώς η ύπαρξη επιχειρηματικών μοντέλων προϋποθέτει την ικανοποίηση των χρηστών και επομένως την εκπλήρωση και διασφάλιση των απαιτήσεων που αυτοί θέτουν και εκφράζονται ως απαιτήσεις ποιότητας (γίνεται αναλυτική αναφορά στην κατηγοριοποίηση των απαιτήσεων στην ενότητα 5.4 της διατριβής)

Η σελίδα αυτή είναι σκόπιμα λευκή

Επιπρόσθετα και με στόχο την κάλυψη μέρους των παραπάνω ερευνητικών προτεραιοτήτων και απαιτήσεων, αναλύεται και προτείνεται μια καινοτόμος αρχιτεκτονική πλέγματος επόμενης γενιάς στην ενότητα 4.4.

1.3 Συνεισφορά – Καινοτομία Διατριβής

Η διατριβή αυτή παρουσιάζει μια σειρά ζητημάτων που αποτελούν συνεισφορά και καινοτομία στον ευρύτερο χώρο των πλεγμάτων. Αρχικά και αναφορικά με τις ροές εργασίας, γίνεται λεπτομερής ανάλυση των συγκεκριμένων κατηγοριών αυτών και της γλώσσας περιγραφής OWL-WS [99] ως επέκταση της OWL-S [15] για την περιγραφή των διεργασιών μιας Ροής Εργασίας και του καθορισμού συγκεκριμένων ροών (Concrete Workflows) που περιέχουν τόσο σημασιολογικές πληροφορίες όσο και πληροφορίες εκτέλεσης της ροής.

Στη συνέχεια παρουσιάζεται μια καινοτόμος αρχιτεκτονική πλέγματος επόμενης γενιάς, που χρησιμοποιώντας τις θεμελιώδεις έννοιες της αντικειμενοστραφούς αρχιτεκτονικής και της OGSA, διαμορφώνει το πλαίσιο μέσα στο οποίο είναι εφικτή η περιγραφή και εκτέλεση δυναμικών ροών εργασίας. Η συγκεκριμένη αρχιτεκτονική που παρουσιάστηκε εστιάζεται κυρίως στο μοντέλο εικονικής υποδομής (Grid VIM) το οποίο στοχεύει στην αφαίρεση τόσο των επιχειρησιακών διεργασιών, όσο και των απαραίτητων αλληλεπιδράσεων με την υποδομή πλέγματος με τη χρήση προσαρμοσμένων ροών εργασίας (Adaptive Workflows) και τη δυναμική δέσμευσης υπηρεσιών που περιλαμβάνει σημασιολογική εύρεση και επιλογή υπηρεσιών. Η λογική της εφαρμογής μπορεί να αναπαρασταθεί με αφηρημένες ροές εργασίας, οι οποίες περιλαμβάνουν λειτουργικούς περιορισμούς της εφαρμογής. Κατά τη διάρκεια της εκτέλεσης μιας ροής, αφηρημένες εργασίες της εφαρμογής μετατρέπονται σε

συγκεκριμένες για τις οποίες είναι γνωστός ο τρόπος εκτέλεσής τους σε περιβάλλον πλέγματος.

Η κύρια συμβολή αυτής της εργασίας, ωστόσο, βρίσκεται στο μηχανισμό απεικόνισης δυναμικών ροών εργασίας με στόχο την παροχή και διασφάλισης της προσφερόμενης Ποιότητας Υπηρεσιών (Quality of Service - QoS). Αρχικά επιχειρείται μια ενδελεχή περιγραφή των συστατικών στοιχείων και του τρόπου λειτουργίας του μηχανισμού και στη συνέχεια γίνεται ανάλυση αυτού και σύγκριση με τις σύγχρονες τάσεις που συναντώνται στην παγκόσμια βιβλιογραφία. Η βασική λειτουργικότητα του μηχανισμού αναφέρεται στην επιλογή υπηρεσιών για την εκτέλεση των διεργασιών που περιγράφονται σε μια δεδομένη ροή εργασίας. Η προαναφερθείσα επιλογή γίνεται λαμβάνοντας υπόψη τις παραμέτρους ποιότητας που καθορίζουν ένα επιθυμητό επίπεδο ποιότητας υπηρεσιών από τη πλευρά του χρήστη και ένα αντίστοιχο επίπεδο παροχής από την πλευρά των φορέων παροχής υπηρεσιών σε υποδομές πλέγματος. Στο σημείο αυτό θα πρέπει να τονιστεί ότι ενώ η αρχική επιλογή γίνεται για τις επιμέρους διεργασίες της ροής εργασίας, ο μηχανισμός στη συνέχεια επιχειρεί βελτιστοποίηση των επιλογών για το σύνολο της ροής ώστε να επιτύχει συνολική ποιότητα υπηρεσίας (end-to-end QoS). Επιπλέον, μια από τις παραμέτρους είναι ο χρόνος εκτέλεσης της ροής και λόγω αυτού ο μηχανισμός απεικόνισης περιλαμβάνει και ένα σχήμα χρονοπρογραμματισμού για την επίτευξη των χρονικών ορίων ποιότητας. Η διαφορά μεταξύ συστημάτων που συναντώνται στη βιβλιογραφία και του συγκεκριμένου μηχανισμού έγκειται στο γεγονός ότι τα συστήματα αυτά αναφέρονται στην επιλογή υπηρεσιών και κόμβων με βάση τις παραμέτρους ποιότητας αντιμετωπίζοντας μόνο τις ειδικές περιπτώσεις ελαχιστοποίησης μιας από τις παραμέτρους ενώ επιπλέον όλες οι παράμετροι έχουν τον ίδιο συντελεστή βαρύτητας. Στο μηχανισμό που αναπτύχθηκε οι παράμετροι υπηρεσίας αντιμετωπίζονται και με συνδυασμένο τρόπο ενώ επίσης εισάγεται η περίπτωση όπου ο χρήστης ορίζει προτιμήσεις και, επομένως, μια από τις παραμέτρους μπορεί να παίζει

σημαντικότερο ρόλο κατά τη διαδικασία επιλογής υπηρεσιών και λόγω αυτού της εκχωρείται ένας συντελεστής βαρύτητας.

Τέλος, σημειώνεται ότι η παροχή δυνατοτήτων απεικόνισης ροών εργασίας με βάση μεγέθη ποιότητας είναι υψηλής σημασίας για τις επιχειρήσεις καθώς επιτρέπει την υιοθέτηση διαφόρων επιχειρηματικών μοντέλων με σχεδιασμό της ροής εργασίας σύμφωνα με τα μεγέθη ποιότητας και εκ των προτέρων γνώση του επιπέδου παροχής αυτής για τους χρήστες. Επίσης, ο συγκεκριμένος μηχανισμός απεικόνισης ροής εργασίας επιτρέπει την επιλογή των υποστάσεων υπηρεσιών στα πλαίσια δεδομένων ροών εργασίας με βάση την πληροφορία ποιότητας προκειμένου να ικανοποιούνται οι απαιτήσεις των χρηστών. Η υιοθέτηση διαφορετικών επιχειρηματικών μοντέλων και στόχων επιτυγχάνεται με αξιολόγηση των εναλλακτικών στρατηγικών και παρακολούθηση των μεταβολών που ενδεχομένως να συμβούν και έχουν σημαντική επίδραση στις στρατηγικές, στις μεθοδολογίες και στη δομή των επιχειρηματικών διαδικασιών. Επομένως, η εκτέλεση μιας ροής εργασίας σύμφωνα με τις αρχικές προδιαγραφές ποιότητας μπορεί να απαιτεί εκ νέου χρονοπρογραμματισμό λόγω μη αναμενόμενων εξελίξεων, καθυστερήσεων ή παραβιάσεων. Όταν απαιτείται προσαρμογή, δημιουργείται ένα σύνολο πιθανών εναλλακτικών, με στόχο την αλλαγή της ροής εργασίας καθώς η ποιότητα θα πρέπει να συνεχίζει να πληροί τις αρχικές απαιτήσεις. Για κάθε εναλλακτική υπηρεσία, πριν την υιοθέτησή της σε μια ροή εργασίας, είναι απαραίτητο να εκτιμηθεί η επίδρασή της στη συνολική παρεχόμενη ποιότητα της ροής εργασίας. Αυτό επιτυγχάνεται όχι μόνο με επιλογή των βέλτιστων υποστάσεων υπηρεσιών σε δεδομένη ροή εργασίας αλλά με χαρακτηρισμό των μη επιλεγμένων υποστάσεων προκειμένου να είναι εκ των προτέρων γνωστό αν χρειάζονται σε μια διαδικασία εκ νέου ο χρονοπρογραμματισμός της. Κάθε φορά που ολοκληρώνεται μια αλλαγή σε μια ροή εργασίας, γίνεται εκ νέου εκτίμηση του επιπέδου της παρεχόμενης ποιότητας υπηρεσιών προκειμένου να ικανοποιούνται οι περιορισμοί του χρήστη.

1.4 Οργάνωση της Διατριβής

Το παρόν έγγραφο αποτελείται από οκτώ (8) κεφάλαια. Στις ενότητες των κεφαλαίων αυτών παρουσιάζεται ουσιαστικά και με αναλυτικό τρόπο το αντικείμενο της διδακτορικής διατριβής.

Στο Κεφάλαιο 2 γίνεται μια εισαγωγή στις έννοιες και στις Βασικές Αρχές του πλέγματος, στους όρους, στα χαρακτηριστικά, στην ταξινόμηση και στα πλεονεκτήματα που προκύπτουν από τη χρήση της συγκεκριμένης τεχνολογίας σε διάφορους τομείς.

Στο Κεφάλαιο 3 δίνεται μια αναλυτική περιγραφή των σύγχρονων προσεγγίσεων αναφορικά με τις Δυναμικές Ροές Εργασίας, τον ορισμό και τις κατηγορίες αυτών καθώς επίσης και της γλώσσας που χρησιμοποιείται για την περιγραφή τους σε περιβάλλον πλέγματος.

Στο Κεφάλαιο 4 συζητούνται στοιχεία Αρχιτεκτονικής Σχεδίασης του πλέγματος. Γίνεται αναφορά στις παραμέτρους σχεδίασης και το κεφάλαιο συνεχίζει περιγράφοντας την Αρχιτεκτονική Ανοιχτών Υπηρεσιών Πλέγματος (OGSA) και την Υπηρεσιοστρεφή Αρχιτεκτονική (SOA). Στη συνέχεια περιγράφονται αναλυτικά τα αρχιτεκτονικά στοιχεία πλέγματος που επιτρέπουν την εκτέλεση δυναμικών ροών εργασίας.

Στο Κεφάλαιο 5 παρουσιάζεται η έννοια της Ποιότητας Υπηρεσίας καθώς και η ανάγκη διασφάλισης αυτής για την υιοθέτηση επιχειρηματικών μοντέλων σε περιβάλλον πλέγματος. Επιπλέον, περιγράφονται οι παράμετροι που καθορίζουν και «ποσοτικοποιούν» την ποιότητα και οι κατηγορίες απαιτήσεων ποιότητας για συγκεκριμένες παραμέτρους. Το συγκεκριμένο κεφάλαιο καταλήγει αναφέροντας μοντέλα δρομολόγησης εργασιών για την εκτίμηση και διασφάλιση της προσφερόμενης ποιότητας.

Στο Κεφάλαιο 6 περιγράφεται το μοντέλο και ο Μηχανισμός Απεικόνισης Δυναμικών Ροών Εργασίας σε υπηρεσίες πλέγματος με στόχο την παροχή ποιότητας υπηρεσιών. Προτείνεται ένας αλγόριθμος για την προαναφερθείσα απεικόνιση και ένα σχήμα καθορισμού

περισσότερων συγκεκριμένων ροών εργασίας. Επίσης, περιγράφεται η ενσωμάτωση του συγκεκριμένου μηχανισμού στην αρχιτεκτονική που παρουσιάστηκε στο τέταρτο κεφάλαιο.

Στο Κεφάλαιο 7 γίνεται αξιολόγηση της λειτουργίας και της απόδοσης του μηχανισμού απεικόνισης δυναμικών ροών εργασίας και επιλογής υπηρεσιών καθώς και της αποτελεσματικότητάς του με χρήση ενός σεναρίου σε περιβάλλον πλέγματος για την εφαρμογή επεξεργασίας τρισδιάστατων εικόνων (3D Rendering). Επίσης παρουσιάζεται το σύστημα δοκιμών και τρεις (3) προσομοιώσεις με τα αντίστοιχα αποτελέσματα.

Τέλος, στο Κεφάλαιο 8 περιλαμβάνεται η σύνοψη της διατριβής και τα συμπεράσματα που εξήχθησαν κατά την εκπόνησή της καθώς και η κύρια συνεισφορά της στο χώρο της τεχνολογίας πλέγματος και συζητούνται θέματα μελλοντικής εργασίας και επέκτασης των ερευνητικών αποτελεσμάτων.

Η σελίδα αυτή είναι σκόπιμα λευκή

2

Γενικά για το περιβάλλον

Πλέγματος

Η χρήση ηλεκτρονικών υπολογιστών σε περιβάλλον πλέγματος [2] 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 [3], [4]. Στο κεφάλαιο αυτό παρουσιάζονται θέματα για το περιβάλλον πλέγματος όπως ιστορικά στοιχεία, όροι, ταξινόμηση και πλεονεκτήματα από τη χρήση αυτού.

2.1 Ιστορικά στοιχεία - από το Διαδίκτυο στο Πλέγμα

Η έρευνα για την τεχνολογία πλέγματος ξεκίνησε στους ακαδημαϊκούς χώρους έστω και στα πλαίσια απλών συζητήσεων, πολλά χρόνια πριν ο επιχειρησιακός κόσμος αντιληφθεί τις δυνατότητες, που θα του πρόσφερε η μετάβαση σε ένα σύστημα κατανεμημένων υπολογισμών. Όμως οι αρχικές ιδέες, που σχετίζονταν με τον όρο πλέγμα, 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.

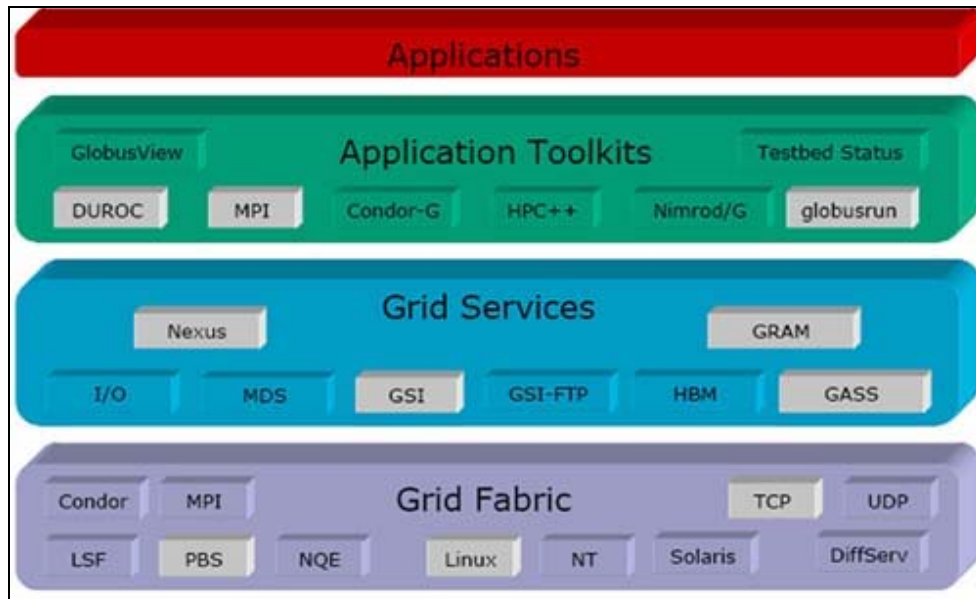
Η αρχιτεκτονική των cluster συστημάτων παρουσίαζε ορισμένα βασικά προβλήματα στην περίπτωση της επικοινωνίας μεταξύ clusters διαφορετικών ιδρυμάτων:

- 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 υπηρεσιών πλέγματος.

Η ανάπτυξη δικτύων υψηλών ταχυτήτων και η ανάγκη για μεγάλη επεξεργαστική ισχύ οδήγησε σε έντονη ερευνητική δραστηριότητα στον τομέα της τεχνολογίας πλέγματος. 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 στο σχήμα της επόμενης σελίδας (Σχήμα 2).



Σχήμα 2: Αρχιτεκτονική Globus

Η συνέχεια γίνεται με την ανάπτυξη των Web Services [2001] , υπηρεσιών που είναι προσβάσιμες μέσω του διαδικτύου και χρησιμοποιούν συγκεκριμένα πρωτόκολλα περιγραφής (όπως XML, SOAP, WSDL) και τέλος με την εγκαθίδρυση της Αρχιτεκτονικής Ανοιχτών Υπηρεσιών Πλέγματος (Open Grid Service Architecture - OGSA) [2002] ως κύριας αρχιτεκτονικής της τεχνολογίας πλέγματος (όπως περιγράφεται στην ενότητα 4.3). Το πρότυπο αυτό είναι το πιο σημαντικό για τις υποδομές πλέγματος, αφού περιγράφει τις δυνατότητες του συστήματος αυτού να αναλύσει την λειτουργία του σε όλα τα επίπεδά του.

2.2 Χαρακτηριστικά

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

Η σελίδα αυτή είναι σκόπιμα λευκή

- Επιτρέπουν το διαμοιρασμό των πόρων σε πολλαπλούς χρήστες διαφορετικών κοινοτήτων με ετερογενή πεδία εφαρμογών και γεωγραφική κατανομή. Ένα Grid 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 (για παράδειγμα 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 πλέγματος 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 διαμοιρασμό 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 [49], 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 (για παράδειγμα 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 πλέγματα 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 πλέγματα 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 διαδικτύου, έχουν ξεκινήσει οι προσπάθειες για την 3η γενιά Grids (**3G Grids**), που βασίζονται στην Αρχιτεκτονική

Ανοιχτών Υπηρεσιών Grid (όπως περιγράφεται στην ενότητα 4.3), όπου μια σειρά από προδιαγραφές κοινών και ανοιχτών διεπαφών υποστηρίζουν τη διαλειτουργικότητα ανεξάρτητα ανεπτυγμένων υπηρεσιών. Η πρόσφατα εκδοθείσα προδιαγραφή Open Grid Services Infrastructure - OGSi) είναι ο θεμέλιος λίθος της παραπάνω αρχιτεκτονικής. Με την εισαγωγή προτυποποιημένων τεχνικών προδιαγραφών, η 3η γενιά πλέγματος θα επιταχύνει τον ανταγωνισμό και την επίτευξη διαλειτουργικότητας όχι μόνο μεταξύ εφαρμογών και εργαλειαθικών, αλλά κυρίως μεταξύ διαφορετικών υλοποιήσεων βασικών υπηρεσιών πλέγματος.

2.4 Όροι Πλέγματος

2.4.1 To 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 πλέγματος. 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 πλέγματος 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 η τεχνολογία πλέγματος 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 απαλλαχθούν 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 και του WSRF (WS-Resource Framework – βλ. ενότητα 4.3.2.5). Το τελευταίο χρησιμοποιείται για τη δημιουργία υπηρεσιών δικτύου οι οποίες προσφέρουν στους μετέχοντες σε περιβάλλον πλέγματος 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 πλέγματος είναι ο ελεγχόμενος και συντονισμένος διαμοιρασμός και χρήση πόρων για την επίλυση προβλημάτων στο πλαίσιο δυναμικών *Εικονικών Οργανισμών*, E.O. (Virtual Organizations-VOs) [3]. Ο παραπάνω διαμοιρασμός αφορά όχι μόνο στην ανταλλαγή δεδομένων αλλά επίσης στην άμεση πρόσβαση σε οντότητες πλέγματος (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 γίνουν μέλη ή να αποχωρούν 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 E.O. ταυτόχρονα με αποτέλεσμα να υπάρχει δυνατότητα οι VO να επικαλύπτονται μεταξύ τους.

Βάσει των παραπάνω, υπάρχουν πολλές απαιτήσεις για την επιθυμητή λειτουργία των E.O., 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 E.O. 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 (πόρος2).

- 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 παροχείς 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.

2.4.6 Συμβόλαια Παροχής Υπηρεσιών

2.4.6.1 Γενικά

Τα Συμβόλαια Παροχής Υπηρεσιών (Service Level Agreement - SLA) [53], [55] χρησιμοποιούνται για την κωδικοποίηση των ιδιοτήτων υπηρεσιών και τη συμφωνία των εμπλεκόμενων μερών (παρόχου και καταναλωτή). Οι προαναφερθείσες ιδιότητες εκφράζονται ως όροι ενός συμβολαίου, οι οποίοι μπορούν να χωριστούν στους υψηλού

επιπέδου (συνχά νομικούς όρους) και τους χαμηλού επιπέδου(συνήθως τεχνικούς όρους). Οι τεχνικοί όροι μπορούν να υποβληθούν σε επεξεργασία αυτόματα όταν τίθενται σε ένα κατάλληλο σχήμα. Περιγράφουν χαρακτηριστικά της παρεχόμενης υπηρεσίας και της ποιότητα των υπηρεσιών που έχουν συμφωνηθεί με τους όρους υψηλού επιπέδου.

Τα συμβόλαια παροχής υπηρεσιών καθορίζουν ένα σύνολο μη λειτουργικής ποιότητας των ιδιοτήτων υπηρεσιών που είναι σημαντικές τόσο για τους παρόχους όσο και για τους καταναλωτές και εκφράζουν την πρόθεση της μελλοντικής παροχής υπηρεσιών (μπορούν επομένως να χρησιμοποιηθούν για να προβλέψουν τη μελλοντική χρησιμοποίηση των πόρων του πλέγματος). Οι φορείς παροχής υπηρεσιών μπορούν να χρησιμοποιήσουν αυτό το στοιχείο για τον προγραμματισμό και τη δέσμευση των πόρων στις υπηρεσίες καθώς και για τον υπολογισμό της αξίας των προσφερόμενων υπηρεσιών.

2.4.6.2 Διαπραγμάτευση

Η συμφωνία ενός συμβολαίου περιλαμβάνει μια φάση διαπραγμάτευσης κατά την οποία ο πάροχος και ο καταναλωτής διαπραγματεύονται στους όρους του συμβολαίου. Αμφότερα τα συμβαλλόμενα μέρη έχουν τους στόχους και τις προτεραιότητές τους σχετικά με τους προαναφερθέντες όρους. Οι καταναλωτές απαιτούν συγκεκριμένα επίπεδα ποιότητας της προσφερόμενης υπηρεσίας ενώ οι φορείς παροχής υπηρεσιών επιθυμούν τη χρήση των ανενεργών πόρων τους με ένα βέλτιστο τρόπο. Η έκβαση της φάσης διαπραγμάτευσης είναι είτε μια συμφωνία (το συμβόλαιο SLA) είτε η λήξη της διαπραγμάτευσης. Στην τελευταία περίπτωση, ο καταναλωτής υπηρεσιών πρέπει να βρει έναν άλλο πάροχο που μπορεί να ικανοποιήσει τις απαιτήσεις του.

Καθώς ένας καταναλωτής μπορεί να είναι σε στάδιο διαπραγμάτευσης με περισσότερων του ενός προμηθευτές, αντίστοιχα και ένας προμηθευτής μπορεί να έχει τρέχουσες διαπραγματεύσεις με πολλούς καταναλωτές. Και στις δυο παραπάνω περιπτώσεις το άλλο

συμβαλλόμενο μέρος δεν ξέρει για τις άλλες αλληλεπιδράσεις και λόγω αυτού η διαπραγμάτευση θεωρείται πάντα ως διμερής αλληλεπίδραση.

2.4.6.3 Συμβιβασμός Όρων

Οι καταναλωτές και οι προμηθευτές έχουν διαφορετικούς στόχους κατά τη διαπραγμάτευση SLAs [53]. Οι καταναλωτές απαιτούν συγκεκριμένες ιδιότητες ποιότητας υπηρεσίας (για παράδειγμα εκτέλεση της εργασίας μέχρι μια ορισμένη προθεσμία) ενώ οι φορείς παροχής υπηρεσιών επιθυμούν τη μεγιστοποίηση των εσόδων τους. Αυτοί οι διαφορετικοί στόχοι μπορούν να συμβιβαστούν με τη συμφωνία στα συμβόλαια. Ο καταναλωτικός στόχος μπορεί να εκφραστεί υπό μορφή απαιτήσεων με όρους SLA. Οι επιχειρησιακοί στόχοι του προμηθευτή θέτουν αντίστοιχα ορισμένους περιορισμούς με όρους SLA. Αν και οι απαιτήσεις και οι περιορισμοί φαίνονται να είναι διαφορετικοί, μπορούν και οι δύο να εκφραστούν με τον ίδιο τρόπο.

Για τις αριθμητικές τιμές, η κατάσταση μπορεί εύκολα να εκφραστεί στους μαθηματικούς όρους. Έστω ότι, με βάση το παραπάνω παράδειγμα, ο καταναλωτής έχει την απαίτηση για ένα ανώτερο όριο (μέγιστο) a σε έναν όρο x και ότι ο πάροχος μπορεί να εξασφαλίσει ένα ελάχιστο όριο b για τον ίδιο όρο. Τότε:

$$\text{Καταναλωτής:} \quad x \leq a \quad (1)$$

$$\text{Πάροχος:} \quad b \leq x \quad (2)$$

$$\text{Συμβιβασμός:} \quad b \leq x \leq a, \text{ εάν } b \leq a \quad (3)$$

$$0, \text{ εάν } b > a \quad (4)$$

Το αποτέλεσμα, δηλαδή η σειρά των έγκυρων τιμών για το x όπου επιτυγχάνεται συμβιβασμός είναι το διάστημα $[a : b]$. Εάν $a < b$, ο συμβιβασμός παράγει το κενό σύνολο και οι περιορισμοί των δύο συμβαλλόμενων μερών δεν μπορούν να ικανοποιηθούν

ταυτόχρονα. Αν και οι παραπάνω εξισώσεις λειτουργούν με τα απέραντα διαστήματα, η ίδια λογική ισχύει και όταν οι δύο άκρες των διαστημάτων είναι κλειστές.

2.4.6.4 Μέρη ενός Συμβολαίου

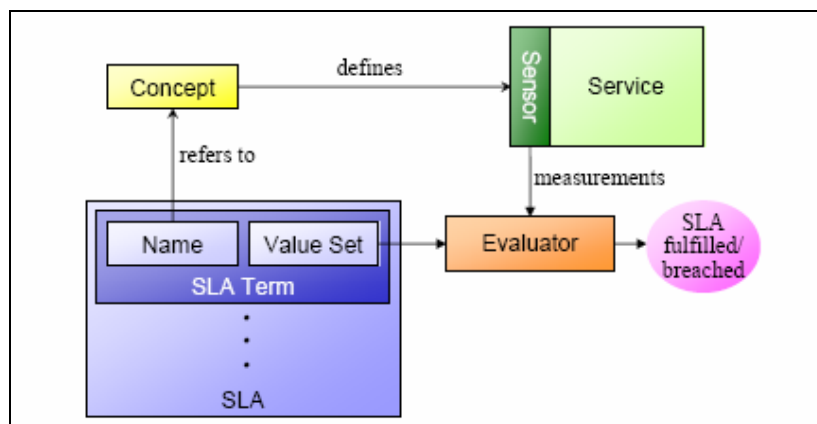
Σε ένα συμβόλαιο SLA μπορούν να προσδιοριστούν διάφορα τμήματα, όπως το γενικότερο πλαίσιο και οι όροι. Το πλαίσιο περιέχει στις ιδιαίτερες πληροφορίες για τους συμμετέχοντες του SLA και τα στοιχεία πληρωμής. Οι όροι καθορίζουν συλλογικά τις ιδιότητες (και επομένως την ποιότητα) της υπηρεσίας που παρέχεται μέσα από αυτό το συμβόλαιο.

Οι πληροφορίες που περιλαμβάνονται στο γενικότερο πλαίσιο δεν χρησιμοποιούνται για την απόφαση αποδοχής ή απόρριψης μιας πρότασης σύμβασης SLA. Η μόνη εξαίρεση είναι τα στοιχεία πληρωμής και ειδικότερα η τιμή υπηρεσιών που διαδραματίζουν επίσης ένα ρόλο στην απόφαση αποδοχής / απόρριψης. Πρέπει να σημειωθεί ότι η τιμή είναι μόνο μέρος του τμήματος στοιχείων πληρωμής του γενικότερου πλαισίου. Άλλες παράμετροι είναι οι μορφές πληρωμής (για παράδειγμα πιστωτικές κάρτες) ή οι ποινικές ρήτρες. Για να αξιολογηθούν και αυτές οι πληροφορίες θα πρέπει να είναι μέρος του τμήματος όρων.

Οι όροι που δηλώνονται σε ένα πρότυπο συμβόλαιο SLA ή μια προσφορά χρησιμοποιούνται για να υπολογιστεί ο βαθμός συμμόρφωσης των όρων με τις απαιτήσεις του πελάτη. Ο βαθμός συμμόρφωσης μπορεί να χρησιμοποιηθεί για την ταξινόμηση των προτύπων / προσφορών ή την απόρριψη μερικών ή όλων. Ανάλογα με το πρωτόκολλο διαπραγμάτευσης, δύναται είτε η αντικατάσταση των όρων που δεν ικανοποιούν τις απαιτήσεις του πελάτη με όρους που τον ικανοποιούν και επανέναρξη της διαπραγμάτευσης με το φορέα παροχής υπηρεσιών για να συμφωνήσουν με την τροποποιημένη προσφορά είτε η απόρριψη των προτύπων που δεν ικανοποιούν τις απαιτήσεις του πελάτη και η επιλογή του καταλληλότερου (σύμφωνα με μερικές μετρικές) μεταξύ των υπολοίπων.

2.4.6.5 Όροι Συμβολαίων

Οι όροι των συμβολαίων SLA εκφράζουν ορισμένες παραμέτρους της υπηρεσίας που παρέχεται και δεν αναφέρονται στην περιγραφή της λειτουργίας που παρέχεται (δηλώνεται κάπου αλλού και δεν ανήκει στους όρους). Όλοι οι συμμετέχοντες πρέπει να γνωρίζουν σε ποια έννοια ένας όρος αναφέρεται και καθώς ο καθορισμός μιας έννοιας μπορεί να είναι αρκετά σύνθετος, ένας όρος συμβολαίου SLA περιέχει συνήθως και ένα προσδιοριστικό για την προαναφερθείσα έννοια.



Σχήμα 3: Σχέσεις όρων σε συμβόλαιο παροχής υπηρεσιών

Όπως φαίνεται στο παραπάνω σχήμα (Σχήμα 3), σε έναν όρο η έννοια συσχετίζεται με ένα σύνολο τιμών που είναι αποδεκτές στο συμβόλαιο. Κατά τη διάρκεια ισχύος του συμβολαίου πραγματοποιούνται μετρήσεις, οι οποίες πρέπει να περιλαμβάνονται στο σύνολο αποδεκτών τιμών μιας έννοιας. Εάν μια μέτρηση είναι εκτός των έγκυρων τιμών, το SLA παραβιάζεται (SLA Violation) και ο παραβιάζοντας συμμετέχων υπόκειται σε μια πληρωμή ποινικής ρήτρας.

Το σύνολο έγκυρων τιμών που διευκρινίζονται στους όρους SLA είναι συνήθως μια σειρά που διευκρινίζει ένα ελάχιστο ή ένα μέγιστο όριο για κάποια υπηρεσίας. Ένα παράδειγμα είναι ο χρόνος απόκρισης, για το οποίο ένα SLA θα περιείχε ένα ανώτερο. Δεδομένου ότι οποιαδήποτε αξία κάτω από εκείνο το μέγιστο είναι αποδεκτή, ο όρος στο SLA θα

Η σελίδα αυτή είναι σκόπιμα λευκή

διευκρίνιζε πιθανότατα τη σειρά από 0 μέχρι το μέγιστο όριο ως έγκυρη σειρά των τιμών για το χρόνο απόκρισης. Ένα παρόμοιο παράδειγμα είναι το διαθέσιμο εύρος ζώνης σε κάποια υπηρεσία για το οποίο ένας όρος SLA θα διευκρίνιζε ένα ελάχιστο εύρος ζώνης που πρέπει πάντα να είναι διαθέσιμο. Δεδομένου ότι ένα μεγαλύτερο εύρος ζώνης θα ήταν επίσης αποδεκτό, ο όρος θα διευκρίνιζε τη σειρά από το χαμηλότερο όριο στο άπειρο ως αποδεκτές τιμές για το εύρος ζώνης.

Τα παραπάνω εκφράζονται ως εξής με μαθηματικούς όρους:

$$\text{Χρόνος Απόκρισης } t: \quad t < t_{\max} \text{ όπου } t \in [0..t_{\max}] \quad (5)$$

$$\text{Εύρος Ζώνης } b: \quad b > b_{\min} \text{ όπου } b \in [b_{\min}..\infty) \quad (6)$$

Οι καταναλωτές υπηρεσιών και οι φορείς παροχής έχουν διαφορετικούς όρους για τις διάφορες παραμέτρους. Η διαφορά αναφέρεται στο όριο (μέγιστο ή ελάχιστο) των παραμέτρων. Στο ανωτέρω παράδειγμα, οι φορείς παροχής υπηρεσιών θα προτιμούσαν τα υψηλότερα όρια στο χρόνο απόκρισης και τα χαμηλότερα στο εύρος ζώνης ενώ αντίθετα οι καταναλωτές θα προτιμούσαν τα χαμηλότερα όρια στο χρόνο απόκρισης και τα υψηλότερα όρια στο εύρος ζώνης. Κατά τη διάρκεια της φάσης διαπραγμάτευσης ενός SLA, αυτές οι διαφορές πρέπει να καμφθούν ώστε να υπάρξει συμφωνία και συμβόλαιο SLA. Αυτός ο συμβιβασμός αποτελεί εργασία της υπηρεσίας του διαμεσολαβητή όπως αναφέρεται στην ενότητα 4.4.2.6 της παρούσης διατριβής.

2.5 Ταξινόμηση Υποδομών Πλέγματος

Υπάρχουν διάφορες ταξινομήσεις των υποδομών 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 τους. Στην περίπτωση αίτησης χρήσης πόρων από δύο ή και περισσότεροι 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 2.4.5). Ένας Ε.Ο. μπορεί να περιλαμβάνει πολλά partner grids και ένα partner grid πολλούς Ε.Ο.. 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 περιβάλλον πλέγματος 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 μία από τις υπάρχουσες υποδομές πλέγματος και αποκτώντας σε αντάλλαγμα οικονομικό κέρδος. Το open Grid (συχνά αναφέρεται και ως παγκόσμιο πλέγμα) 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 Σύγκριση Τεχνολογίας Πλέγματος με άλλες Τεχνολογίες

Ο σχεδιασμός του πλέγματος 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 4.3).

Η δομή του πλέγματος θεωρούμε ότι αποτελεί εξέλιξη άλλων δομών και τεχνολογιών, όπως τα κατανεμημένα συστήματα (distributed systems) και συγκεκριμένα τα υπολογιστικά περιβάλλοντα μεγάλης κλίμακας [88], 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 υποδομών πλέγματος παρουσιάζει πολλές ομοιότητες με άλλες τεχνολογίες (στις οποίες βασίστηκε) αλλά και πολλές αντιθέσεις. Αναλυτικότερα:

- Όπως και στο Διαδίκτυο τα πλέγματα κρύβουν όλη την πολυπλοκότητα, και οι χρήστες απολαμβάνουν μια απλή και υλοποιημένη υπηρεσία. Σε αντίθεση όμως με το Διαδίκτυο το οποίο απλά χρησιμεύει για τη διευκόλυνση της επικοινωνίας, τα περιβάλλοντα πλέγματος επιτρέπουν πλήρη συνεργασία, με στόχο την επίτευξη κοινών επιχειρηματικών στόχων.
- Όπως και τα P2P συστήματα οι υποδομές πλέγματος 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 κατανεμημένα συστήματα (ή clusters) τα πλέγματα ενοποιούν υπολογιστικούς πόρους. Σε αντίθεση με τα κατανεμημένα συστήματα (ή clusters) τα οποία χρειάζονται γεωγραφική εγγύτητα και ομοιογένεια των συστημάτων, τα συστήματα πλέγματος μπορεί να είναι γεωγραφικά κατανεμημένα και ετερογενή.
- Όπως και οι τεχνολογίες εικονικής οργάνωσης συστημάτων (virtualization) τα πλέγματα 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 πλέγματος επιτρέπουν την εικονική οργάνωση απέραντων, ανόμοιων, απομακρυσμένων πόρων.

2.7 Πλεονεκτήματα

Πολλές εταιρίες επιδιώκουν να αξιοποιήσουν τα πλεονεκτήματα της σημερινής δομής του πλέγματος 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 πλέγματος, αλλά και να υιοθετήσουν καινούργιες καθώς αυτές εξελίσσονται, προσφέροντας νέα πλεονεκτήματα και ευκολίες.

Η πλατφόρμα του πλέγματος μπορεί να συνδυάζει τους περισσευούμενους πόρους σε ένα εταιρικό δίκτυο δημιουργώντας ένα ισχυρό περιβάλλον 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 πλέγματος, χωρίς να περιορίζονται από μελλοντικές εξελίξεις και προοπτικές.

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

- Χαμηλότερο κόστος υπολογισμού, καθώς η σχέση απόδοσης / τιμή είναι διαφορετική. Οι διάφορες υποδομές πλέγματος, 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 περιβάλλον πλέγματος 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 το περιβάλλον πλέγματος, 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.

2.8 Τομείς που ευνοούνται από την τεχνολογία Πλέγματος

Η τεχνολογία πλέγματος 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 πλέγματος 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 πλέγματος 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 πλέγματος. Τράπεζες, ασφαλιστικές εταιρείες, χρηματιστηριακές επιχειρήσεις και αναλυτές, όλοι έχουν τη δυνατότητα να βελτιώσουν την ποιότητα των υπηρεσιών που παρέχουν κάνοντας χρήση της τεχνολογίας πλέγματος, η οποία μπορεί να παρέχει γρήγορες 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 τεχνολογίας πλέγματος. Εδώ υπάρχουν δύο κατηγορίες χρηστών. 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 τεχνολογίας πλέγματος. Η δεύτερη κατηγορία, που σχετίζεται κυρίως με τη δημόσια έρευνα, θεωρείται κυρίως ως καταναλωτής της τεχνολογίας πλέγματος. 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 τεχνολογίας πλέγματος, 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 τεχνολογίας πλέγματος μπορεί να αποδειχθεί εξαιρετικά ωφέλιμη. Στη 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

Δυναμικές Ροές Εργασίας

Όπως έχει ήδη αναφερθεί, με βάση διεισδυτικά πρότυπα διαδικτύου, τα πλέγματα επιτρέπουν την κοινή χρήση υπολογιστικών και πληροφοριακών πόρων εκτός των ορίων επιμέρους τμημάτων ή οργανισμών, με ασφαλή και εξαιρετικά αποτελεσματικό τρόπο. Τα πλέγματα υποστηρίζουν την κοινή χρήση, τη διασύνδεση και τη χρήση διαφόρων πόρων, ενσωματωμένων στο πλαίσιο ενός δυναμικού υπολογιστικού συστήματος.

Η διαχείριση των λειτουργιών της ροής εργασίας κάποιας εφαρμογής σε περιβάλλον πλέγματος απαιτεί την ενορχήστρωση των προαναφερθέντων κατανεμημένων πόρων [5]. Σε αυτό το πλαίσιο, η ροή εργασίας είναι σημαντικός παράγοντας για τη σύνθεση των εφαρμογών σε πλέγματα, προσφέροντας αρκετά πλεονεκτήματα [6], όπως:

- Δυνατότητα ανάπτυξης δυναμικών εφαρμογών με ενορχήστρωση κατανεμημένων πόρων
- Αξιοποίηση πόρων που βρίσκονται σε συγκεκριμένο τομέα με στόχο την αύξηση της απόδοσης ή τη μείωση των δαπανών εκτέλεσης
- Εκτέλεση πολλών απομακρυσμένων διαχειριστικών τομέων με στόχο την απόκτηση συγκεκριμένων δυνατοτήτων επεξεργασίας

- Ενοποίηση πολλών ομάδων που συμμετέχουν στη διαχείριση διαφορετικών τμημάτων της ροής εργασίας, ενισχύοντας έτσι τις συνεργασίες μεταξύ οργανισμών

3.1 Ορισμοί

Καθώς η ροή εργασίας είναι μια ευρεία τεχνολογική έννοια στη συγκεκριμένη παράγραφο διευκρινίζεται η ορολογία που αφορά τους ορισμούς της ροής εργασίας που χρησιμοποιούνται στη συνέχεια της παρούσης διατριβής.

Όσο αφορά το γενικό ορισμό, η ομάδα WfMC (Workflow Management Coalition) δίνει τον εξής ορισμό [7]: *"Ροή εργασίας είναι η αυτοματοποίηση μιας επιχειρηματικής διαδικασίας, στο σύνολό της ή εν μέρει, κατά την οποία έγγραφα, πληροφορίες ή εργασίες μεταφέρονται μεταξύ των συμμετεχόντων προκειμένου να εκτελεστεί κάποια δράση, σύμφωνα με ένα σύνολο διαδικαστικών κανόνων."*

Διαφορετικά, η ροή εργασίας μπορεί να οριστεί ως η ενορχήστρωση ενός συνόλου δραστηριοτήτων με στόχο την επίτευξη ενός πολύπλοκου στόχου, ενώ στα περιβάλλοντα πλέγματος, η ροή εργασίας περιλαμβάνει διαδικασίες εφαρμογών, επιχειρηματικές διαδικασίες και διαδικασίες υποδομής [8]. Η ροή εργασίας είναι ένας αρχιτεκτονικά σημαντικός παράγοντας για τη δυναμική διαλειτουργικότητα και προσαρμογή σε διαφορετικά επιχειρηματικά μοντέλα, που μπορούν να συμπεριληφθούν στη γενική έννοια πολιτικές ροής εργασίας, και πλαίσια ανάπτυξης. Τα μοντέλα ροής εργασίας χρησιμοποιούνται προκειμένου να οριστεί μια ροή εργασίας τόσο σε επίπεδο εργασιών όσο και σε επίπεδο δομής.

Η έρευνα που διεξάγεται για την τεχνολογία πλέγματος έχει διατυπώσει ακόμη δύο ορισμούς, που είναι προσανατολισμένοι κυρίως προς την επιστήμη των υπολογιστών:

Workflow είναι ένα σχέδιο αλληλεπίδρασης διάφορων επιχειρηματικών διεργασιών, χωρίς να αντιστοιχεί απαραίτητα σε ένα καθορισμένο σύνολο από επιχειρηματικές διεργασίες. Όλες αυτές οι αλληλεπιδράσεις μπορεί να είναι μεταξύ υπηρεσιών (services) που λειτουργούν σε ένα μόνο κέντρο δεδομένων ή μεταξύ ενός εύρους από διαφορετικές πλατφόρμες και υλοποιήσεις οπουδήποτε.

Grid Workflow είναι ένας εύκολος τρόπος δημιουργίας νέων υπηρεσιών συνδυάζοντας ήδη υπάρχουσες υπηρεσίες. Μια νέα υπηρεσία (service) μπορεί να δημιουργηθεί και να χρησιμοποιηθεί καταχωρώντας τον ορισμό ενός workflow σε μία workflow engine.

Στη συνέχεια και βάσει των παραπάνω ορισμών παρουσιάζονται οι κατηγορίες των ροών εργασίας.

3.2 Κατηγορίες

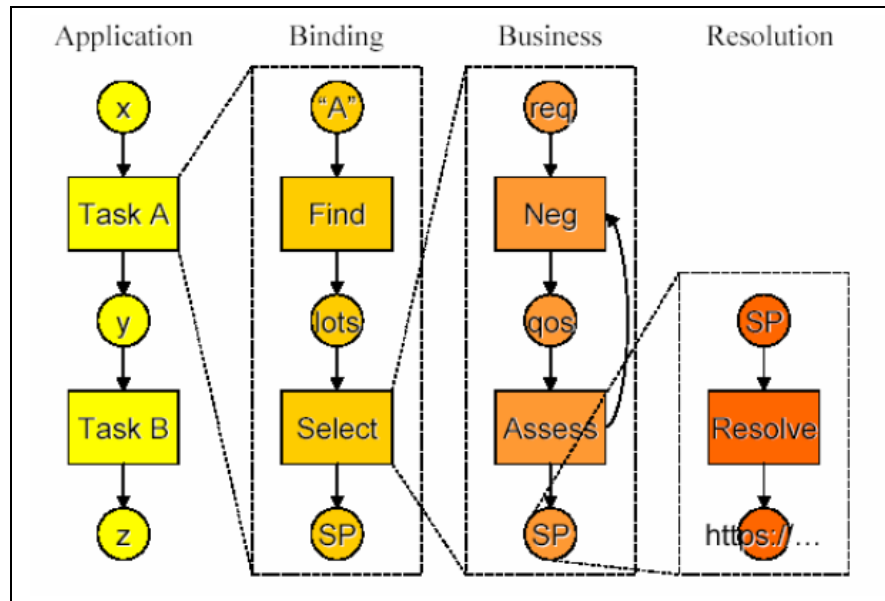
Οι ροές εργασίας παίζουν πολύ σημαντικό ρόλο στην αρχιτεκτονική του πλέγματος κυρίως όσον αφορά την ευέλικτη και δυναμική συνάθροιση των υπηρεσιών που προσφέρονται μέσω του πλέγματος. Χρησιμεύει στην αναπαράσταση και σύνθεση της λειτουργικής και επιχειρηματικής λογικής σε μια εφαρμογή.

Ανάλογα με το περιεχόμενο των ροών εργασίας και αναφορικά με τη περιγραφή αυτού, μπορεί να είναι γενικές, ώστε να περιλαμβάνουν την περιγραφή των βημάτων μιας εφαρμογής (application workflow) είτε πιο συγκεκριμένα ώστε να περιλαμβάνουν και να περιγράφουν τις διεργασίες που αφορούν όχι καθαυτή την εκτέλεση της εφαρμογής, αλλά τη χρήση υπηρεσιών, που αφορούν την εύρεση και επιλογή των υπηρεσιών, που υλοποιούν κάθε βήμα της εφαρμογής. Βάσει αυτού, η εκτέλεση κάθε βήματος στη ροή εργασίας μιας εφαρμογής έχει ως αποτέλεσμα τη δυναμική δημιουργία επιπλέον ροών εργασίας. Για παράδειγμα, η εκτέλεση ενός βήματος (Task A) της ροής εργασίας εφαρμογής (application

workflow) έχει ως αποτέλεσμα τη δημιουργία μιας επιπλέον (Binding) ροής εργασίας, που αφορά το συσχετισμό του βήματος με την εύρεση των υπηρεσιών και την επιλογή αυτής, που τελικά θα χρησιμοποιηθεί για την υλοποίηση του βήματος. Στο δεύτερο αυτό workflow, το βήμα της επιλογής της υπηρεσίας εισάγει μιας νέας (Business) ροής εργασίας, που περιγράφει τη διαδικασία επιλογής εφαρμόζοντας τις παραμέτρους ποιότητας υπηρεσίας. Στο τέλος της επιλογής εισάγεται και μια τελευταία (Resolution) ροή εργασίας, που περιγράφει τα βήματα που γίνονται προκειμένου να χρησιμοποιηθεί η επιλεγθείσα υπηρεσία για την υλοποίηση του Task A. Από αυτό το σχήμα διαπιστώνουμε τελικά ότι όχι μόνο η εφαρμογή αλλά και οι επιχειρησιακές (business) διεργασίες που χρησιμοποιούνται από το σύστημα, εισάγουν workflows που μπορούν να αναπαρασταθούν και να διαχειρισθούν με έναν ενιαίο τρόπο από το σύστημα.. Δομικά αυτά οι ροές εργασίας δε διαφέρουν, αλλά διαφέρουν σίγουρα ως προς τις σημασιολογικές πληροφορίες που τα συνοδεύουν. Με βάση αυτή την παρατήρηση μπορούμε να χωρίσουμε τις ροές εργασίας σε δύο κατηγορίες:

- Ροές Εργασίας Εφαρμογών (application workflows), που περιλαμβάνουν workflows που αναπτύσσονται και διαχειρίζονται από τους δημιουργούς της εφαρμογής, έτσι ώστε να μπορούν να εκτελεστούν σε περιβάλλον πλέγματος. Δηλαδή ο δημιουργός της εφαρμογής ή γενικά όποιος θέλει να την εκτελέσει σε υποδομή πλέγματος δημιουργεί μια ροή εργασίας εφαρμογής, που υλοποιεί την εργασία που θέλει να πραγματοποιήσει.
- Ροές Εργασίας Πολιτικής (policy workflows), που περιλαμβάνουν επιχειρησιακές διεργασίες, που αναπαρίστανται και αναπτύσσονται ώστε να επιτρέπουν τη δυναμική αναπαράσταση της ροής εργασίας.

Τα παραπάνω παρουσιάζεται στο σχήμα της επόμενης σελίδας (Σχήμα 4).



Σχήμα 4: Ανάλυση ροών εργασίας από την ροή της εφαρμογής

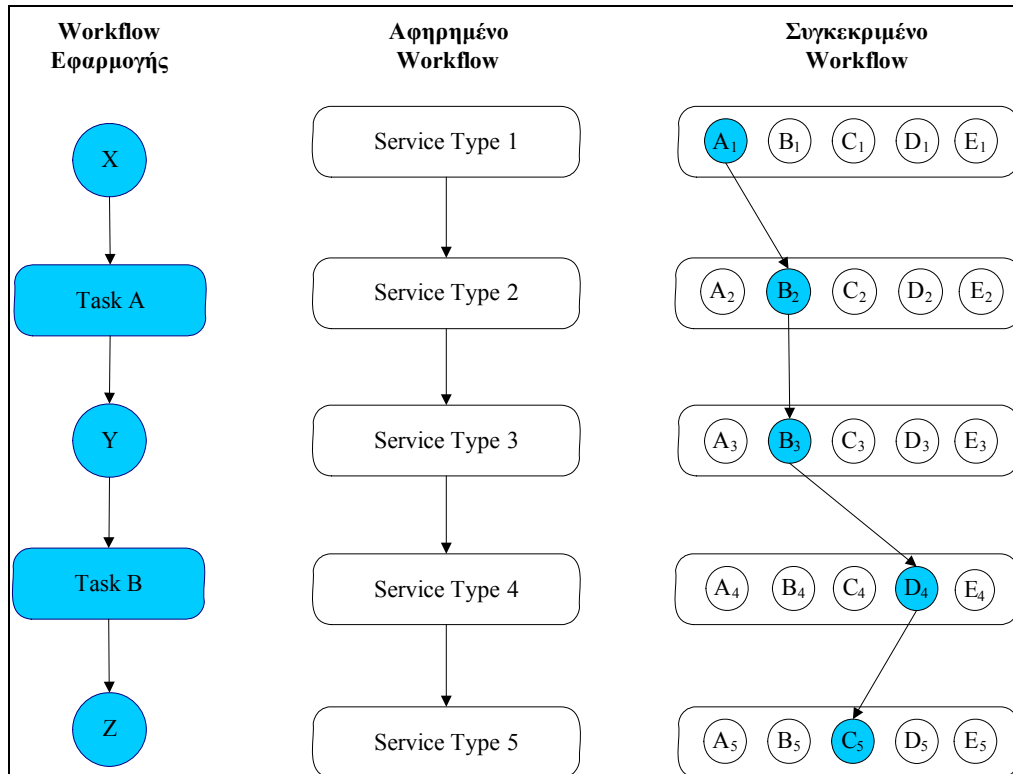
Επιπρόσθετα οι ροές εργασίας χωρίζονται και ανάλογα με το αν είναι γνωστό πως ακριβώς θα υλοποιηθούν οι διεργασίες που αυτά περιλαμβάνουν. Σε αυτή την περίπτωση έχουμε τις ακόλουθες δύο κατηγορίες - *αφηρημένες (abstract)* και *συγκεκριμένες (concrete)* [9], [10]:

- Συγκεκριμένη ροή εργασίας (Concrete workflow) - αναφέρονται επίσης και ως εκτελέσιμες ροές εργασίας στη βιβλιογραφία [11] - ονομάζεται μία σύνθεση υπηρεσιών, η οποία παρέχει τόσο σημασιολογικές πληροφορίες όσο και πληροφορίες εκτέλεσης, οι οποίες αφορούν τη σύνθεση του workflow, τόσο από την πλευρά των υπηρεσιών, που το αποτελούν όσο και από την πλευρά της πλήρης σύνθεσής του. Ένα συγκεκριμένο workflow μπορεί να θεωρηθεί ως ένα μονοπάτι, που προκύπτει από την επιλογή στιγμιότυπων των υπηρεσιών έτσι ώστε να επιτευχθεί το workflow της εφαρμογής (application workflow).
- Αφηρημένη ροή εργασίας (abstract workflow). Για να κατασκευαστεί ένα συγκεκριμένο workflow, συνδυάζεται το workflow εφαρμογής (application workflow) με τις παραμέτρους ποιότητας υπηρεσίας (QoS). Ένα ενδιάμεσο βήμα προς το στόχο αυτό αποτελεί η κατασκευή ενός αφηρημένου workflow, το οποίο

Η σελίδα αυτή είναι σκόπιμα λευκή

προκύπτει ως η σύνθεση των υπηρεσιών που υλοποιούν την εφαρμογή. Παρέχονται όμως μόνο πληροφορίες για το είδος της υπηρεσίας που χρειάζεται σε κάθε βήμα καθώς και οι είσοδοι και έξοδοι αυτού. Δεν περιλαμβάνει πληροφορίες για το ποιος κόμβος συγκεκριμένα θα είναι αυτός που τελικά θα παρέχει την υπηρεσία.

Οι προαναφερθείσες κατηγορίες παρουσιάζονται στο σχήμα που ακολουθεί (Σχήμα 5):



Σχήμα 5: Ορισμοί ροής εργασίας

Αναλύοντας περισσότερο τις παραπάνω κατηγορίες, στα αφηρημένα μοντέλα, οι εργασίες περιγράφονται με μη συγκεκριμένη μορφή χωρίς να γίνεται αναφορά σε συγκεκριμένους πόρους πλέγματος για την εκτέλεση των εργασιών αυτών ενώ παράλληλα δίνεται η δυνατότητα στους χρήστες να ορίσουν ροές εργασίας με ευέλικτο τρόπο, απομονώνοντας τις λεπτομέρειες εκτέλεσης. Επίσης, τα αφηρημένα μοντέλα παρέχουν μόνο πληροφορίες για την έννοια των υπηρεσιών και σχετικά με τον τρόπο σύνθεσης της ροής εργασίας. Επομένως, είναι εφικτή η κοινή χρήση των περιγραφών της ροής εργασίας μεταξύ των

Η σελίδα αυτή είναι σκόπιμα λευκή

χρηστών του πλέγματος, γεγονός που είναι κρίσιμης σημασίας για τους συμμετέχοντες σε εικονικούς οργανισμούς (βλ. ενότητα 2.4.5).

Στα συγκεκριμένα μοντέλα, οι εργασίες της ροής εργασίας αντιστοιχίζονται σε συγκεκριμένους πόρους και, επομένως, το μοντέλο παρέχει πληροφορίες σχετικά με την έννοια των υπηρεσιών και την εκτέλεση σχετικά με τον τρόπο σύνθεσης της ροής εργασίας, τόσο για τις υποστάσεις υπηρεσιών όσο και για τη συνολική σύνθεση (για παράδειγμα σχέσεις στη ροή δεδομένων, δομές ροής ελέγχου). Σε αντιστοιχία με τα αφηρημένα μοντέλα και σε σχέση με τους εικονικούς οργανισμούς, οι εργασίες που περιλαμβάνονται σε ένα συγκεκριμένο μοντέλο μπορεί επίσης να αναφέρονται σε αιτήματα μετακίνησης δεδομένων με στόχο τη δημοσιοποίηση των προσφάτως εξαγμένων δεδομένων στον εικονικό οργανισμό [9]. Πρέπει να διευκρινιστεί ότι οι υποστάσεις υπηρεσιών δεν αντιστοιχούν απαραίτητα σε πόρους του πλέγματος εφόσον, εντός κάποιου πόρου, μπορεί να υπάρχουν και να εκτελούνται περισσότερες από μια υποστάσεις υπηρεσιών.

Με βάση τους ορισμούς των μοντέλων ροής εργασίας, οι εργασίες στα αφηρημένα μοντέλα είναι φορητές και μπορούν να απεικονισθούν σε οποιαδήποτε κατάλληλη υπηρεσία πλέγματος με χρήση κατάλληλων μηχανισμών ανίχνευσης και απεικόνισης. Σε αυτή την κατεύθυνση, ένα από τα θέματα της παρούσης διατριβής είναι ένας νέος μηχανισμός απεικόνισης ροής εργασίας, το αποτέλεσμα του οποίου είναι μια συγκεκριμένη ροή εργασίας. Η συγκεκριμένη ροή εργασίας μπορεί να αντιμετωπιστεί ως "διαδρομή", μια επιλογή υποστάσεων υπηρεσιών από τύπους υπηρεσιών με συγκεκριμένη σειρά προκειμένου να επιτευχθεί ο στόχος της εκτέλεσης της ροής εργασίας μιας συγκεκριμένης εφαρμογής.

3.3 Καθορισμός Συγκεκριμένης Ροής Εργασίας

Όπως αναφέρθηκε στην προηγούμενη παράγραφο, ένα από τα αντικείμενα της παρούσας διδακτορικής διατριβής είναι ο προσδιορισμός και η περιγραφή της διαδικασίας που πρέπει να εφαρμοσθεί προκειμένου να οριστεί η συγκεκριμένη ροή εργασίας με δεδομένα τη ροή εργασίας μιας εφαρμογής και τις απαραίτητες παραμέτρους ποιότητας υπηρεσιών (QoS). Αυτό το θέμα έρευνας αναφέρεται ως *Περιορισμοί QoS Ροής Εργασίας* (Workflow QoS Constraints) και παραμένει ένας από τους σημαντικούς παράγοντες των Συστημάτων Διαχείρισης Ροής Εργασίας σε Πλέγμα (Grid Workflow Management System) και, ειδικότερα, στο πεδίο Σχεδιασμού Ροής Εργασίας (Workflow Design) [12].

Γενικά, οι μηχανισμοί απεικόνισης ροής εργασίας είναι αναπόσπαστα τμήματα της παροχής ποιότητας υπηρεσιών καθώς αυτός είναι ο μόνος τρόπος για να εκτιμηθεί, υπολογιστεί και καθοριστεί η τελική απεικόνιση των ροών εργασίας καθώς και η επιλογή των διαθέσιμων τύπων και υποστάσεων υπηρεσιών προκειμένου να προσφερθεί ποιότητα υπηρεσιών στον τελικό χρήστη από ένα σύνολο παρόχων υπηρεσιών.

3.4 Γλώσσα Περιγραφής

Η δυναμική φύση των ροών εργασίας αναφέρεται στην προσαρμογής τους και στη δέσμευση στο χρόνο εκτέλεσης επιχειρησιακών διεργασιών σύμφωνα με διάφορα μοντέλα. Ο δημιουργός της ροής εργασίας της εφαρμογής δε χρειάζεται να είναι ενήμερος για το πώς ακριβώς θα εκτελεστεί. Αρκεί να παρέχονται μαζί με την περιγραφή της ροής και στοιχεία για το πώς θα έπρεπε να εκτελεστεί (όπως για παράδειγμα ποιες είναι οι είσοδοι κάθε διεργασίας, τι εξόδους δίνουν, με ποιες άλλες επικοινωνούν κ.α.). Ο τρόπος που παρέχονται αυτές οι πληροφορίες θα πρέπει να είναι τυποποιημένος, έτσι ώστε να υπάρχει ένας κοινός τρόπος περιγραφής. Επίσης χρειάζεται ένας ενιαίος τρόπος περιγραφής των υπηρεσιών που

παρέχονται, το τι κάνουν, από ποιες διεργασίες αποτελούνται, ποιος τις παρέχει και ποιες είναι οι προϋποθέσεις για να τις χρησιμοποιήσει κάποιος. Πιο συγκεκριμένες απαιτήσεις που πρέπει να πληροί μια γλώσσα περιγραφής ροών εργασίας είναι οι ακόλουθες:

- Δυνατότητα ορισμού αφηρημένων (abstract) ροών εργασίας. Πρέπει να είναι δυνατή η περιγραφή ροών χωρίς να καθορίζονται συνδέσεις μεταξύ κάθε εργασίας και μιας υπηρεσίας που την υλοποιεί, έτσι ώστε οι συνδέσεις να γίνονται στο χρόνο εκτέλεσης.
- Σημασιολογική περιγραφή εργασιών. Κάθε εργασία σε μια αφηρημένη ροή εργασίας πρέπει να διαθέτει μια περιγραφή ή τουλάχιστον να είναι ενός καθορισμένου τύπου εργασίας, επιτρέποντας έτσι την εύρεση κατά το χρόνο εκτέλεσης μια υπηρεσίας που να την υλοποιεί.
- Απεικόνιση ροών εργασίας. Πρέπει να υπάρχει η δυνατότητα δημιουργίας συνδέσεων αφηρημένων εργασιών με πιο λεπτομερείς ροές, οι οποίες θα εισάγονται στη θέση των εργασιών κατά το χρόνο εκτέλεσης.
- Διαχείριση ροών εργασίας υψηλής τάξης. Πρέπει να υπάρχει η δυνατότητα να αντιμετωπίζονται οι ροές εργασίας ως δεδομένα, έτσι ώστε μια εργασία να μπορεί να έχει ως είσοδο μια ροή εργασίας και να επιστρέφει μια άλλη ως έξοδο.

Έχουν αναπτυχθεί διάφορες γλώσσες περιγραφής ροών εργασίας όπως για παράδειγμα η BPEL [14] και η OWL-S [15]. Η γλώσσα που δημιουργήθηκε είναι η OWL-WS και καλύπτει την ανάγκη ενιαίας περιγραφής για όλες τις προηγούμενες απαιτήσεις αποτελεί επέκταση της σημασιολογικής γλώσσας OWL-S, που χρησιμοποιείται για την περιγραφή υπηρεσιών.

3.5 Η γλώσσα *OWL-WS*

3.5.1 Χρήσιμοι ορισμοί

Στην ενότητα 4.2.2 περιγράφονται αναλυτικά οι Διαδικτυακές Υπηρεσίες (Web Services). Στο σημείο αυτό και για λόγους κατανόησης της παρούσης ενότητας συνοπτικά αναφέρεται ότι αποτελούν λογισμικές οντότητες σχεδιασμένες για να υποστηρίξουν την επικοινωνία και την αλληλεπίδραση μεταξύ υπολογιστών ή εφαρμογών στο διαδίκτυο. Διαθέτουν διεπαφές με τη χρήση των οποίων καθίσταται εφικτή η επικοινωνία μεταξύ των υπηρεσιών. Μεταφέροντας αυτόν τον ορισμό σε περιβάλλον πλέγματος μπορεί να ειπωθεί ότι μια υπηρεσία πλέγματος είναι μια υπηρεσία δικτύου, που σχεδιάστηκε για να λειτουργεί σε περιβάλλον πλέγματος και να ικανοποιεί τις απαιτήσεις του πλέγματος στο οποίο συμμετέχει.

Στο σημείο αυτό ο όρος «υπηρεσία πλέγματος» αναφέρεται σε μια υπηρεσία σχεδιασμένη για να λειτουργεί σε περιβάλλον πλέγματος χρησιμοποιώντας επιπλέον και ένα είδος σημασιολογικής πληροφορίας. Μια τέτοια υπηρεσία μπορεί να είναι και ένα εκτελέσιμο πρόγραμμα ή η διαδικασία μεταφοράς δεδομένων, τα οποία διατίθεται μέσω μιας διεπαφής υπηρεσίας. Με την ίδια λογική ακόμη και οι ροές εργασίας μπορούν να θεωρηθούν σύνθετες υπηρεσίες καθώς αποτελούνται από πολλές άλλες υπηρεσίες. Επιπρόσθετα, μια υπηρεσία μπορεί να θεωρηθεί ως μια πολύ απλή ροή εργασίας. Δηλαδή υπηρεσίες και ροές εργασίας μπορούν να θεωρηθούν ως παρόμοιες λειτουργικές οντότητες και να υφίστανται παρόμοια διαχείριση. Λόγω των παραπάνω, όσα αναφέρονται για τις υπηρεσίες θα ισχύουν και για τις ροές εργασίας.

Οι υπηρεσίες και οι ροές εργασίας μπορεί να είναι αφηρημένες (abstract) ή συγκεκριμένες βάσει του ορισμού που δόθηκε στην ενότητα 3.1. Επιπρόσθετα στο σημείο αυτό θα πρέπει να σημειωθεί ότι υπάρχει η δυνατότητα μια ροή εργασίας να είναι αφηρημένη και

συγκεκριμένη, δηλαδή να παρέχονται και πληροφορίες εκτέλεσης για μερικές μόνο υπηρεσίες. Τότε ονομάζεται ενδιάμεση ροή εργασίας («Intermediate Workflow»), αλλά μπορεί να θεωρηθεί ως ένα είδος αφηρημένης ροής.

Όσον αφορά τις σημασιολογικές πληροφορίες αυτές μπορούν να διακριθούν σε δύο κατηγορίες:

- **Περιορισμοί.** Είναι περιγραφή των απαιτήσεων μιας αφηρημένης υπηρεσίας, όσον αφορά τις απαιτούμενες ιδιότητες της υπηρεσίας. Αυτή η περιγραφή παρέχεται συνήθως από τον πελάτη που αναζητά μια υπηρεσία που να διαθέτει τις ιδιότητες αυτές.
- **Δυνατότητες.** Είναι οι πληροφορίες που προσφέρει η περιγραφή μιας υπηρεσίας που παρέχεται και περιγράφουν τις ιδιότητες της προσφερόμενης υπηρεσίας. Αυτές οι πληροφορίες συνήθως παρέχονται από αυτούς που προσφέρουν την υπηρεσία.

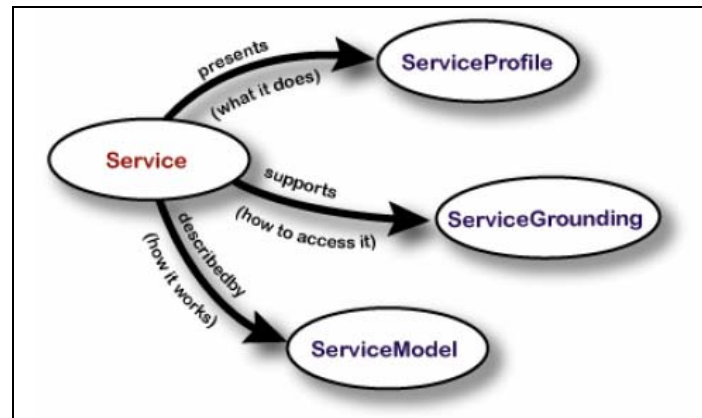
3.5.2 Τα βασικά στοιχεία της γλώσσας OWL-S

Η OWL-S είναι γλώσσα που χρησιμοποιείται για περιγραφή υπηρεσιών. Μερικά από τα βασικά στοιχεία που χρησιμοποιούνται για το σκοπό αυτό είναι:

- Η κλάση «Service», που παρέχει ένα οργανωτικό σημείο αναφοράς για μια δηλωμένη υπηρεσία δικτύου. Για κάθε υπηρεσία που παρέχεται πρέπει να υπάρχει και ένα στιγμιότυπο της κλάσης Service.
- Το προφίλ της υπηρεσίας («Service Profile»), που περιγράφει το τι κάνει η υπηρεσία. Αυτή η μορφή αναπαράστασης εμπεριέχει μια περιγραφή του τι δύναται να πετύχει η υπηρεσία, περιορισμούς στην εφαρμογή της, και στην ποιότητα υπηρεσίας και προϋποθέσεις που ο αιτούμενος την υπηρεσία πρέπει να ικανοποιεί έτσι ώστε να χρησιμοποιήσει την υπηρεσία.

- Το μοντέλο της υπηρεσίας («Service Model»), που παρέχει πληροφορίες στον πελάτη για τον τρόπο χρήσης της υπηρεσίας (σημασιολογικό περιεχόμενο των αιτήσεων, συνθήκες κάτω από τις οποίες ένα συγκεκριμένο αποτέλεσμα προκύπτει και πληροφορίες για την ακολουθία των διεργασιών). Εν συντομία, περιγράφει πώς πρέπει να αιτηθεί η υπηρεσία και τι συμβαίνει όταν αυτή χρησιμοποιείται.
- Η θεμελίωση της υπηρεσίας («Service Grounding»), παρέχει πληροφορίες για τον τρόπο πρόσβασης στην υπηρεσία. Τυπικά καθορίζει ένα πρωτόκολλο επικοινωνίας, τη δομή των μηνυμάτων και άλλες λεπτομέρειες της υπηρεσίας, όπως οι αριθμοί των θυρών που χρησιμοποιούνται για τη σύνδεση. Επίσης καθορίζει, για κάθε σημασιολογικό τύπο εισόδου και εξόδου έναν μη διφορούμενο τρόπο ανταλλαγής δεδομένων του τύπου αυτού με την υπηρεσία.
- Η διεργασία υπηρεσίας («Service Process») καθορίζει τους τρόπους με τους οποίους ένας πελάτης μπορεί να αλληλεπιδράσει με την υπηρεσία. Μια ατομική διεργασία (atomic process) είναι μια διεργασία που δέχεται ένα μήνυμα και έχει ως έξοδο ένα άλλο μήνυμα. Διαθέτει πάντα μια θεμελίωση (grounding). Από την άλλη, οι σύνθετες διεργασίες (composite processes) είναι διεργασίες που δέχονται μηνύματα και τα μεταδίδουν στο εσωτερικό τους (στις διεργασίες που τις αποτελούν). Μπορούν να αποτελούνται από απλές διεργασίες ή και σύνθετες. Η αποσύνθεσή τους μπορεί να προσδιοριστεί χρησιμοποιώντας δομές δεδομένων όπως ακολουθία (sequence) και επιλογή υπό συνθήκη («δομές if-then-else»).

Το σχήμα της επόμενης σελίδας (Σχήμα 6) παρουσιάζει τη σχέση μεταξύ των κυρίων κλάσεων της OWL-S:



Σχήμα 6: Βασικά Στοιχεία της OWL-S

3.5.3 Το μοντέλο της γλώσσας OWL-WS

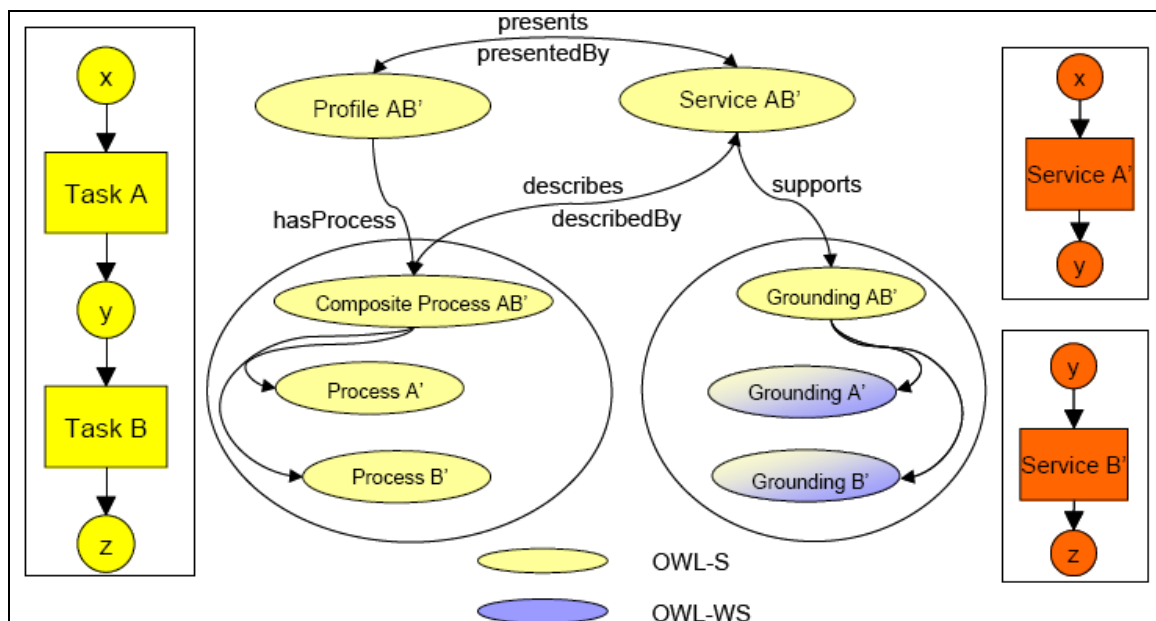
Το μοντέλο της γλώσσας περιγραφής OWL-WS υιοθετεί πλήρως το μοντέλο της OWL-S για να αναπαραστήσει τις συγκεκριμένες υπηρεσίες (Concrete Services). Συγκεκριμένη υπηρεσία είναι μια υπηρεσία με το δικό της προφίλ (που περιέχει μη λειτουργικές πληροφορίες), διεργασία (που περιέχει οδηγίες για την ακολουθία των κλήσεων προς υπηρεσία) και θεμελίωση (που περιέχει αναφορά προς το σημείο που τελικά γίνεται η κλήση). Οι πληροφορίες του προφίλ της αναφέρουν τις δυνατότητες της υπηρεσίας και επομένως είναι σημασιολογικές πληροφορίες που εκφράζουν δυνατότητες.

Ένα επιπλέον στοιχείο είναι ότι η OWL-WS απαιτεί από τα μοντέλα του προφίλ και της διεργασίας να είναι συμβατά μεταξύ τους, κάτι που δεν είναι περιορισμός στη γλώσσα OWL-S.

Το μοντέλο της ροής εργασίας στη OWL-WS βασίζεται στην υπόθεση ότι μια ροή είναι μία σύνθετη υπηρεσία. Έτσι η κάθε ροή διαθέτει δικό της προφίλ, που παρέχει πληροφορίες για τη λειτουργικότητά της. Επίσης ως σύνθεση υπηρεσιών, μπορεί να περιγραφεί από μία σύνθετη διεργασία, που δίνει πληροφορίες για τη δομή της σύνθεσης και τα διάφορα τμήματά της.

Η σελίδα αυτή είναι σκόπιμα λευκή

Όσον αφορά τη συγκεκριμένη ροή εργασίας, αυτή μοντελοποιείται ως μία υπηρεσία (αντικείμενο κλάσης Service), με το δικό της προφίλ, τη δική της σύνθετη διεργασία και μία εκτεταμένη θεμελίωση, της οποίας τα στοιχεία μπορούν να αναφέρονται σε διαφορετική υπηρεσία (που ανήκει στο workflow) η καθεμία. Ενώ στη OWL-S, η διεργασία καθορίζει τα βήματα που απαιτούνται για την αλληλεπίδραση με μια υλοποίηση της υπηρεσίας, στη OWL-WS μια (σύνθετη) διεργασία μπορεί να καθορίζει τα βήματα που απαιτούνται για την αλληλεπίδραση με διάφορες υπηρεσίες έτσι ώστε να υλοποιηθεί η λειτουργικότητα που αναφέρεται στο προφίλ. Αυτό σημαίνει ότι έχουμε πλέον και θεμελίωση (grounding) που αποτελείται από μέρη που αναφέρονται σε διάφορες υπηρεσίες. Ενδεικτικό είναι το σχήμα που ακολουθεί (Σχήμα 7):



Σχήμα 7: Αναπαράσταση συγκεκριμένης ροής εργασίας στις OWL-S, OWL-WS

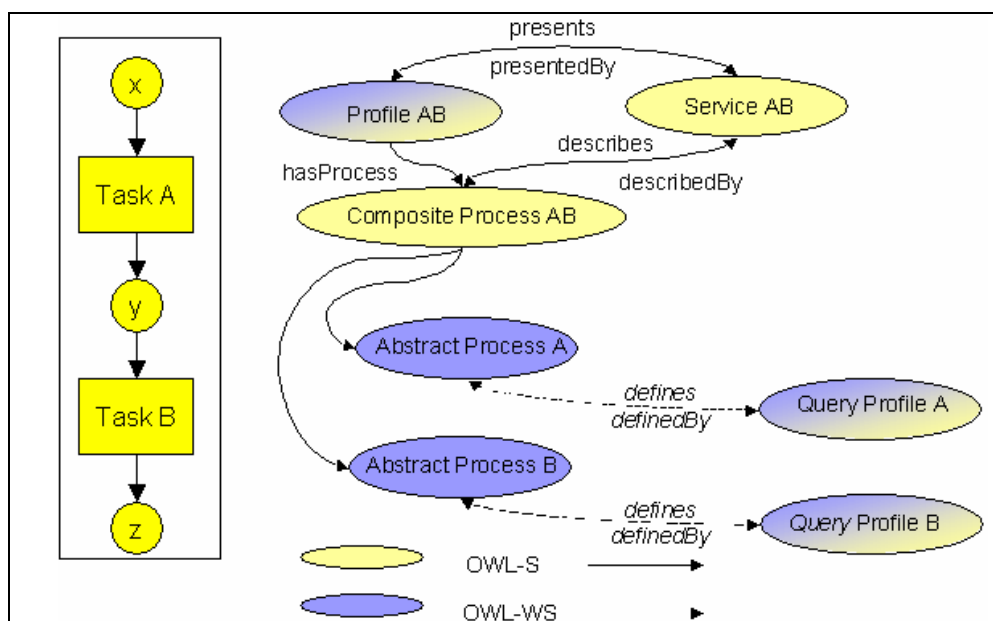
Στο αριστερό μέρος του παραπάνω σχήματος παρουσιάζεται μια αφηρημένη ροή εργασίας και στο δεξιό τα στιγμιότυπα υπηρεσιών που υλοποιούν τις διεργασίες της ροής, ενώ στο κέντρο εμφανίζεται η συγκεκριμένη ροή εργασίας που προκύπτει από το συνδυασμό τους.

Βάσει του σχήματος, η θεμελίωση του συγκεκριμένου workflow είναι σύνθετη και αποτελείται από δύο θεμελιώσεις που αντιστοιχούν σε διαφορετικές υπηρεσίες.

Η σελίδα αυτή είναι σκόπιμα λευκή

Όσον αφορά τις αφηρημένες ροές εργασίας, για να αναπαρασταθούν στη OWL-WS, χρειάζεται να οριστεί η έννοια της αφηρημένης διεργασίας («Abstract Process»). Μία αφηρημένη διεργασία είναι μία ατομική διεργασία για την οποία όμως δεν υπάρχει θεμελίωση, αλλά διαθέτει ένα σχετικό προφίλ. Το προφίλ αυτό περιέχει πληροφορίες για περιορισμούς όσον αφορά την απαιτούμενη λειτουργικότητα της υπηρεσίας που θα υλοποιήσει τη διεργασία. Δηλαδή περιέχει σημασιολογικές πληροφορίες που είναι περιορισμοί. Το προφίλ αυτό λέγεται και προφίλ ερώτησης («Query Profile») καθώς χρησιμοποιείται κατά τη διαδικασία αναζήτησης υπηρεσιών που θα καλύπτουν τις απαιτήσεις της διεργασίας έτσι όπως παρουσιάζονται σε αυτό το προφίλ.

Πλέον μπορεί να οριστεί η αφηρημένη υπηρεσία («Abstract Service») ως μια υπηρεσία (κλάση Service), που διαθέτει μια αφηρημένη διεργασία και το δικό της προφίλ. Έτσι και η αφηρημένη ροή εργασίας θεωρείται ως μια υπηρεσία με το δικό της προφίλ και μια σύνθετη διεργασία, που αποτελείται από αφηρημένες διεργασίες, που καθεμία έχει το δικό της προφίλ. Το σχήμα που αναπαριστά την αφηρημένη ροή εργασίας παρουσιάζεται στην επόμενη σελίδα (Σχήμα 8):



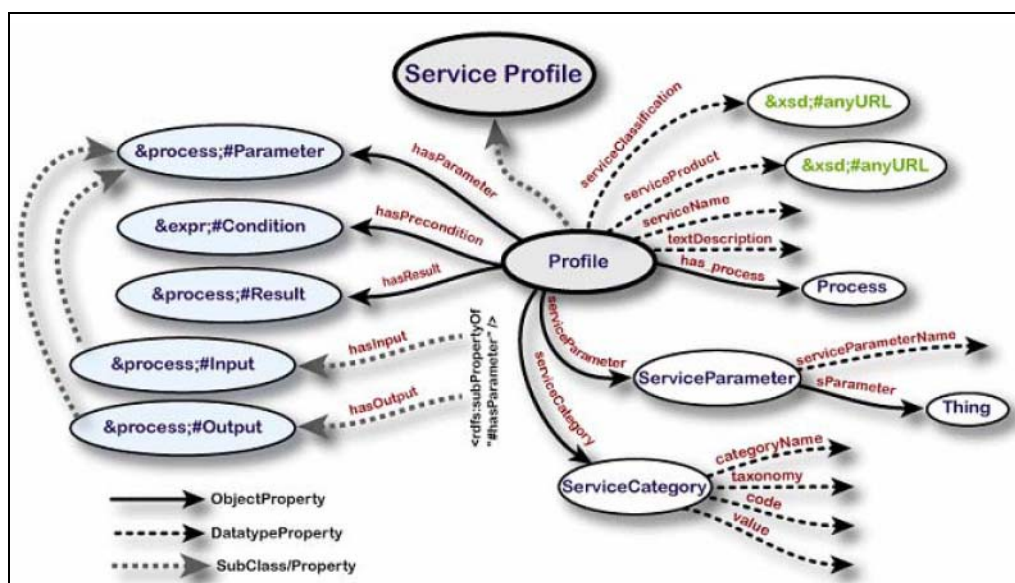
Σχήμα 8: Αναπαράσταση αφηρημένης ροής εργασίας στις OWL-S, OWL-WS

Η σελίδα αυτή είναι σκόπιμα λευκή

Βάσει του προηγούμενου σχήματος, η ροή εργασίας έχει δύο (2) αφηρημένες διεργασίες, που συνοδεύονται από το αντίστοιχο προφίλ. Υπάρχει όμως και το γενικό προφίλ, που πρέπει να είναι συμβατό με τα προφίλ των διεργασιών και το οποίο δύναται να περιέχει επιπλέον πληροφορίες που αφορούν τη ροή εργασίας συνολικά. Τέλος, είναι δυνατό μια ροή εργασίας να περιέχει μια άλλη ροή αυξάνοντας την πολυπλοκότητα των παραπάνω σχημάτων.

3.5.4 Οι παράμετροι των προφίλ

Η γλώσσα OWL-S διαθέτει μηχανισμό για το καθορισμό τόσο των λειτουργικών όσο και των μη λειτουργικών ιδιοτήτων υπηρεσιών. Τα λειτουργικά χαρακτηριστικά μιας υπηρεσίας είναι αυτά που περιγράφουν τι μπορεί να κάνει η υπηρεσία, ενώ τα μη λειτουργικά χαρακτηριστικά της περιγράφουν οτιδήποτε άλλο μπορεί να την αφορά. Αυτά τα χαρακτηριστικά θεωρούνται ως μεταδεδομένα (meta-data) και καθορίζονται μέσα στο προφίλ της υπηρεσίας. Επομένως το προφίλ είναι εκείνο που χαρακτηρίζει την υπηρεσία και το οποίο περιλαμβάνει τελικά όλες τις πληροφορίες αυτής όπως παρουσιάζεται στο παρακάτω σχήμα (Σχήμα 9):



Σχήμα 9: Προφίλ OWL-S

Η σελίδα αυτή είναι σκόπιμα λευκή

Στο αριστερό μέρος του σχήματος παρουσιάζονται οι πιο σημαντικές λειτουργικές ιδιότητες της υπηρεσίας, ενώ στο δεξιό μέρος παρουσιάζονται οι ιδιότητες που χρησιμοποιούνται για να περιγράψουν μη λειτουργικά χαρακτηριστικά. Θα πρέπει να αναφερθεί ότι ο ορισμός των μη λειτουργικών ιδιοτήτων στην OWL-S δεν είναι πλήρης αλλά δίνεται η δυνατότητα ορισμού συγκεκριμένων ιδιοτήτων.

Θα πρέπει στο σημείο αυτό να αναφερθεί ότι για να είναι μια σημασιολογική γλώσσα ροών εργασίας, όπως η OWL-WS αποτελεσματική, είναι βασικό να υπάρχει ένας τρόπος καθορισμού των οντολογιών αναφοράς, που χρησιμοποιούνται για να περιγράψουν τα διάφορα χαρακτηριστικά. Λόγω αυτού οι υπηρεσίες τύπου OWL-S καθορίζουν ένα ειδικό μοντέλο για να προσφέρουν, όποτε είναι δυνατό και χρήσιμο, τη δυνατότητα να ορίζονται πληροφορίες σχετικές με την οντολογία, που θα χρησιμοποιείται για να δίνει μη διφορούμενο νόημα στο χαρακτηριστικό που αναφέρεται.

Στη συνέχεια αναφέρονται οι βασικές ιδιότητες του προφίλ μιας υπηρεσίας στην OWL-WS, που χρησιμοποιούνται για να μοντελοποιήσουν τα λειτουργικά χαρακτηριστικά της. Αυτά τα χαρακτηριστικά προσδιορίζουν την υπηρεσία κατά τη διαδικασία της εύρεσης (discovering), ταιριάζοντας τους λειτουργικούς περιορισμούς, που τίθενται από τον αιτούμενο την υπηρεσία με τις λειτουργικές δυνατότητες, που αναφέρει ο παρέχων την υπηρεσία. Οι ιδιότητες αυτές είναι:

- Κατηγορία υπηρεσίας («Service Category»). Προσφέρει μια ιεραρχία από κατηγορίες και υλοποιείται στην OWL-WS συνδυάζοντας την ιδιότητα «serviceCategory» της OWL-S με την οντολογία.
- Λειτουργικότητα υπηρεσίας («Service Functionality»). Καθορίζει σημασιολογικά τις λειτουργίες μιας υπηρεσίας με βάση τους ειδικούς χειρισμούς που αυτή πραγματοποιεί. Στην OWL-WS υλοποιείται συνδυάζοντας την ιδιότητα

«serviceParameter» της OWL-S με την οντολογία για το συγκεκριμένο χαρακτηριστικό.

- Είσοδοι-έξοδοι υπηρεσίας («Service I/O»). Καθορίζει σημασιολογικά τις δυνατές οντότητες Input-Output-Precondition-Effect (IOPE) που χρησιμοποιούνται. Στην OWL-WS προκύπτει συνδυάζοντας τα στοιχεία IOPE και SWRL (Semantic Web Rule Language). Η SWRL είναι σημασιολογική γλώσσα περιγραφής κανόνων.

Στη συνέχεια ακολουθεί μια λίστα από μη λειτουργικά χαρακτηριστικά του προφίλ μιας υπηρεσίας στη OWL-WS, που σχετίζονται με θέματα όπως ποιότητα υπηρεσίας και ασφάλεια. Στην OWL-S ορίζονται ιδιότητες όπως το όνομα της υπηρεσίας, πληροφορίες επικοινωνίας με τον παρέχοντα την υπηρεσία και άλλες ιδιότητες που χρησιμεύουν στον προσδιορισμό του περιβάλλοντος μέσα στο οποίο παρέχεται η υπηρεσία. Για τη μοντελοποίηση των ιδιοτήτων αυτών στη OWL-WS μπορεί να χρησιμοποιηθεί η ιδιότητα «serviceParameters» που υπάρχει στη OWL-S (είναι μια λίστα παραμέτρων που συνοδεύουν την περιγραφή του προφίλ). Ορισμένες βασικές μη λειτουργικές ιδιότητες είναι οι εξής:

- Ανατομία της υπηρεσίας («Service Anatomy»). Περιλαμβάνει πληροφορίες όπως το όνομα της υπηρεσίας, την περιγραφή της, αν είναι εξαρτώμενη από άλλες υπηρεσίες ή όχι, αν είναι αφηρημένη ή όχι κ.α. Υλοποιείται στην OWL-WS μέσω ιδιοτήτων του προφίλ στη OWL-S, όπως η κατάταξη της υπηρεσίας («serviceClassification»), το όνομα της υπηρεσίας («serviceName») κ.α.
- Διαχείριση υπηρεσίας («Service Management»). Αναφέρεται σε πληροφορίες σχετικά με την επιλογή πολιτικής, όπως «τοπική» (που περιορίζει την ακτίνα δράσης της εύρεσης υπηρεσιών). Υλοποιείται στην OWL-WS συνδυάζοντας την ιδιότητα «serviceParameter» και την αντίστοιχη οντολογία.
- Αξιοπιστία υπηρεσίας («Service Reliability»). Αφορά τον καθορισμό των μηχανισμών χειρισμού λαθών, εξαιρέσεων (exceptions) και λαθών στα συμβόλαια

παροχής υπηρεσιών (βλ. ενότητα 2.4.6). Παραδείγματα ενεργειών είναι η επανάληψη της προσπάθειας (Retry) και η αναστολή της ενέργειας (Suspend). Υλοποιείται στην OWL-WS συνδυάζοντας την ιδιότητα «serviceParameter» και την αντίστοιχη οντολογία.

- Ποιότητα της υπηρεσίας («Quality of Service»). Περιλαμβάνει πληροφορίες για το επίπεδο της ποιότητας, που η υπηρεσία μπορεί να προσφέρει ή απαιτεί. Παραδείγματα είναι περιορισμοί κόστους, χρόνου κ.α. Υλοποιείται στην OWL-WS συνδυάζοντας την ιδιότητα «serviceParameter» και την αντίστοιχη οντολογία.
- Ασφάλεια υπηρεσίας («Service Security»). Καθορίζει τα πρωτόκολλα ασφάλειας τα οποία είναι αποδεκτά, τα απαιτούμενα πιστοποιητικά και τα πρωτόκολλα διαχείρισης παραβίασης της ασφάλειας. Υλοποιείται και αυτή στην OWL-WS συνδυάζοντας την ιδιότητα «serviceParameter» και την αντίστοιχη οντολογία.

Η σελίδα αυτή είναι σκόπιμα λευκή

4

Αρχιτεκτονική Σχεδίαση

Στο συγκεκριμένο κεφάλαιο συζητούνται θέματα σχετικά με την Αρχιτεκτονική Σχεδίαση του πλέγματος. Αρχικά γίνεται αναφορά στις παραμέτρους σχεδίασης και στη συνέχεια περιγράφεται η Αρχιτεκτονική Ανοιχτών Υπηρεσιών Πλέγματος (OGSA) και η Υπηρεσιοστρεφής Αρχιτεκτονική (SOA). Τέλος, το κεφάλαιο περιλαμβάνει ανάλυση και περιγραφή των αρχιτεκτονικών στοιχείων πλέγματος που επιτρέπουν την εκτέλεση δυναμικών ροών εργασίας.

4.1 Παράμετροι Σχεδίασης

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

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 πλέγματος 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 (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.

Ευαισθησία δεδομένων

Στην περίπτωση που τα διαχειριζόμενα δεδομένα είναι ευαίσθητα, το περιβάλλον πλέγματος 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 πλεγμάτων (βλ. ενότητα 2.5), 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 πλέγματος. 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 ευτυχώς ο καθορισμός ασφαλών καναλιών μεταφοράς δεδομένων είναι δυνατός με τη χρήση των προτύπων SSL (Secure Socket Layer) και WSS (Web Services Security). Έτσι είναι δυνατή η κρυπτογράφηση των δεδομένων που μεταφέρονται μεταξύ κόμβων.

Μέγιστες απαιτήσεις σε επεξεργαστική ισχύ

Πριν την εμφάνιση της τεχνολογίας πλέγματος, 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 της τεχνολογίας πλέγματος 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 μέγιστο, να καταφεύγει σε προσφορές επεξεργαστικής ισχύς από τρίτους για να καλύψει τις βραχυπρόθεσμες απαιτήσεις της.

Αποθήκευση δεδομένων

Το περιβάλλον πλέγματος 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 για παράδειγμα 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.

Από τα κριτήρια που αναφέρθηκαν παραπάνω τα σημαντικότερα είναι η χωρητικότητα, ο χρόνος απόκρισης, η ρυθμοαπόδοση και η διάρκεια ζωής των δεδομένων. Με βάση τα χαρακτηριστικά αυτά γίνεται και επιλογή μεταξύ δίσκου ή ταινίας:

- Η χωρητικότητα είναι ο συνολικός χώρος ενός συστήματος αποθήκευσης και θα πρέπει να ανταποκρίνεται στις μέγιστες απαιτήσεις για αποθηκευτικό χώρο. Ορισμένα συστήματα αποθήκευσης δεν κλιμακώνουν αποτελεσματικά, καθώς αυξάνεται η χωρητικότητα. Στη περίπτωση που χρησιμοποιείται δίσκος είναι σήμερα δύσκολο να δημιουργηθεί ένα οικονομικό σύστημα μεγάλης αποθήκευσης για αποθήκευση δεδομένων για μεγάλο χρονικό διάστημα.
- Ο χρόνος απόκρισης συνήθως καθορίζει το μέσο που θα χρησιμοποιηθεί. Όταν το ζητούμενο είναι μικρός χρόνος απόκρισης, τότε είναι αδύνατο να χρησιμοποιηθούν συστήματα βασισμένα σε ταινία, ενώ όταν το ζητούμενο είναι η βέλτιστη απόδοση τότε τα συστήματα ταινίας αποδεικνύονται πιο οικονομικά.
- Το πρόβλημα της ρυθμοαπόδοσης (throughput) αντιμετωπίζεται εύκολα στα περιβάλλοντα πλέγματος. Χάρη στην έμφυτη ικανότητα του πλέγματος να κλιμακώνεται υπάρχει η δυνατότητα να αυξηθεί η ρυθμοαπόδοση χρησιμοποιώντας πολλά στιγμιότυπα του ζητούμενου αντικειμένου, έτσι ώστε να επιτευχθεί αθροιστικά ο ρυθμός παράδοσης των δεδομένων. Επομένως είναι δυνατόν να χρησιμοποιηθούν σχετικά αργά μέσα όπως οι ταινίες αποθήκευσης.
- Τέλος ο χρόνος ζωής των δεδομένων ευνοεί συνήθως τα συστήματα ταινίας

Το σημαντικό στην περίπτωση του πλέγματος είναι ότι με τη χρήση των κατάλληλων πρωτοκόλλων και προτύπων τα διάφορα συστήματα αποθήκευσης αντιμετωπίζονται με ενιαίο τρόπο.

Διαθέσιμο εύρος ζώνης στο διαδίκτυο

Καθώς θεωρείται ότι η τεχνολογία πλέγματος βασίζεται στο διαδίκτυο, το διαθέσιμο εύρος ζώνης (bandwidth) αποτελεί σημαντικό παράγοντα. Επομένως το μέγεθος της σύνδεσης με το internet μιας εφαρμογής που εκτελείται σε περιβάλλον πλέγματος, 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 σε περιβάλλον πλέγματος 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.

Ευτυχώς σήμερα, στην περίπτωση τουλάχιστον των επιστημονικών εφαρμογών, οι συνδέσεις με το διαδίκτυο μπορεί να είναι και της τάξης των μερικών δεκάδων Gigabit/sec. Βέβαια η δυνατότητα μεταφοράς δεδομένων ενός υπολογιστή περιορίζεται στο 1 Gbps, αλλά και μόνο του αυτό σημαίνει δυνατότητα μεταφοράς 16 Gigabyte σε 16 δευτερόλεπτα. 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 των 10 Gbps μπορεί να μην αποδειχθεί αρκετό καθυστερώντας έτσι εν τέλει την εκτέλεση των εργασιών. Επειδή είναι χαρακτηριστικό της υποδομής πλέγματος 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 κατακευκμένου 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.
- Ίσως να 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.

Υπάρχοντες πόροι

Η δυνατότητα της επαναχρησιμοποίησης υπαρχόντων πόρων, συνήθως υπολογιστών, ήταν από τους κύριους λόγους που οδήγησαν στην αποδοχή της τεχνολογίας πλέγματος. Σημαντική 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 μια υποδομή πλέγματος. 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 τεχνολογίας πλέγματος παρέμεναν αναξιόποινα 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 περιβάλλον πλέγματος 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.

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 της τεχνολογίας πλέγματος 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 υπαρχόντων ιδιόκτητων διασυνδέσεων.
- Μια άλλη επιλογή είναι να τρέχει σε ένα μηχάνημα μια υπηρεσία πλέγματος, με την οποία θα επικοινωνεί αυτός που επιθυμεί την πρόσβαση στον πόρο. Αυτή η υπηρεσία πλέγματος θα προωθεί μέσω ειδικού μηχανισμού τις απαραίτητες πληροφορίες στον πόρο. Η δεύτερη αυτή λύση προβλέπεται να είναι φθηνότερη και μάλλον θα προτιμηθεί στο μέλλον.

Μεταφερσιμότητα

Ένα σημαντικό ζήτημα στον αρχιτεκτονικό σχεδιασμό κάθε πλέγματος είναι η μετατροπή των εφαρμογών του προηγούμενου συστήματος έτσι ώστε πλέον να λειτουργούν στο περιβάλλον του πλέγματος και να το αξιοποιούν. Μια ιδέα η οποία αποτελεί γρήγορη και εύκολη λύση είναι να γραφτεί για την εφαρμογή ένας αντίστοιχος *wrapper*, κώδικας ο οποίος θα επιτρέπει τη λειτουργία της στο νέο περιβάλλον και να αξιοποιεί τα ιδιαίτερα

χαρακτηριστικά του. Παρόλα αυτά όμως, για να αξιοποιήσει η εφαρμογή καλύτερα την υποδομή πλέγματος, είναι καλύτερα να προσαρμόζεται ο ίδιος ο κώδικάς της ξαναγράφοντας μέρος του.

Υπάρχουν δύο ειδών συστατικά μέρη στο πλέγμα: αυτά που χρησιμοποιούν τις υπηρεσίες άλλων και αυτά που παρέχουν μια υπηρεσία χωρίς να στηρίζονται σε άλλες υπηρεσίες του πλέγματος. Και στις δύο περιπτώσεις για να δημιουργηθεί ένα τέτοιο στοιχείο, είναι απαραίτητο να λαμβάνονται συνεχώς υπόψη οι χρηστικές και λογισμικές απαιτήσεις του προηγούμενου συστήματος. Για απαιτήσεις, όπως η ασφάλεια, χρειάζεται νέος σχεδιασμός. Από την άλλη ορισμένες απαιτήσεις (όπως απαιτήσεις απόδοσης) γίνονται λιγότερο περιοριστικές στο περιβάλλον πλέγματος (ένας αλγόριθμος ίσως να μη χρειάζεται να είναι υπερβολικά αποδοτικός, αφού οι υπάρχοντες πόροι στο πλέγμα μπορούν να καλύψουν τις απαιτήσεις ενός ίσως μη βελτιστοποιημένου αλγορίθμου).

Ορισμένες υπάρχουσες εφαρμογές μπορούν να εκτελεστούν σε ένα περιβάλλον πλέγματος με τη βοήθεια διεπαφών, που κάνουν χρήση ήδη καθορισμένων προτύπων. Έτσι μειώνεται το φόρτο εργασίας που χρειάζεται, για να γίνουν αυτές προσβάσιμες μέσα σε περιβάλλον πλέγματος. Υπάρχουν όμως και άλλες εφαρμογές, η μεταφορά των οποίων δεν είναι δυνατή ακόμη και με την παραπάνω μέθοδο. Σε αυτή την περίπτωση χρειάζεται για κάθε εφαρμογή να γραφτεί μια ειδική διεπαφή, που θα στηρίζεται στην τεχνολογία πλέγματος.

Δυνατότητες συνεργασίας

Ένα από τα σημαντικότερα ζητήματα κατά το σχεδιασμό υποδομής πλέγματος, είναι οι δυνατότητες που αυτό θα προσφέρει στους χρήστες του για συνεργασία με άλλους χρήστες και οντότητες του συστήματος. Εξάλλου βασικός στόχος του πλέγματος 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.

4.2 Υπηρεσιοστρεφής Αρχιτεκτονική

4.2.1 Γενικά για τις υπηρεσίες

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 της υπηρεσιοστρεφούς αρχιτεκτονικής. 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. Με το σχεδιασμό αυτό είναι εφικτή η δημιουργία ανεξάρτητων μονάδων, ελαστικά συνδεδεμένων μεταξύ τους με κάποιο

πρότυπο πλαίσιο επικοινωνίας. Λόγω αυτής της ανεξαρτησίας των υπηρεσιών, 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.

4.2.2 Διαδικτυακές Υπηρεσίες

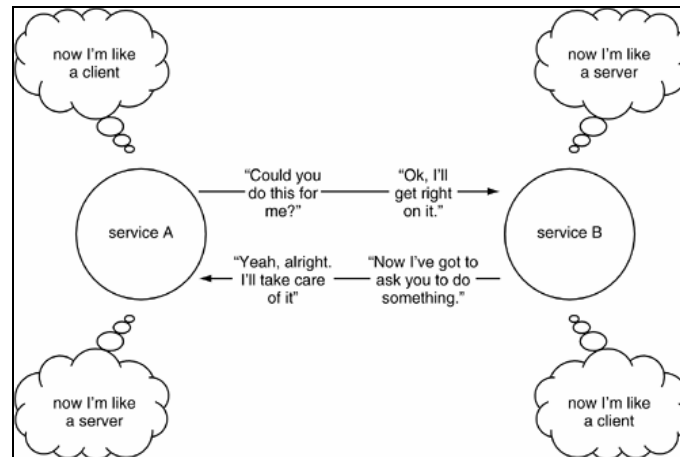
Ο πιο ευρέως διαδεδομένος και επιτυχημένος τύπος υπηρεσιών είναι οι διαδικτυακές υπηρεσίες XML Services γνωστές ως Web Services. Αυτός ο τύπος υπηρεσίας έχει δύο βασικές προαπαιτήσεις:

- Επικοινωνεί μέσω πρωτοκόλλου Internet (κυρίως HTTP)
- Αποστέλλει και δέχεται δεδομένα μέσα από XML αρχεία

Η ευρεία αποδοχή του μοντέλου των διαδικτυακών υπηρεσιών είχε ως αποτέλεσμα την ανάγκη πρόσθετων τεχνολογιών που βασίζονται σε αυτές και τη δημιουργία καινούργιων προτύπων που απαιτεί τα ακόλουθα:

- Περιγραφή της υπηρεσίας αναλυτικά (βλ. ενότητα 4.2.4)
- Δυνατότητα μεταφοράς ενός XML εγγράφου χρησιμοποιώντας SOAP μέσω HTTP (βλ. ενότητα 4.2.5).

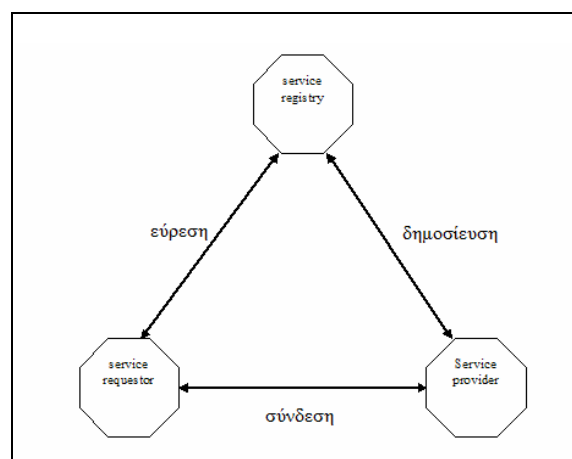
Επιπρόσθετα μια υπηρεσία δύναται να λειτουργεί και ως πελάτης και ως πάροχος υπηρεσίας. Στο σχήμα της επόμενης σελίδας φαίνεται ένα παράδειγμα αυτής της συμπεριφοράς (Σχήμα 10).



Σχήμα 10: Ανταλλαγή ρόλων μεταξύ διαδικτυακών υπηρεσιών κατά τη διάρκεια επικοινωνίας

4.2.3 Πρότυπο Υπηρεσιοστρεφούς Αρχιτεκτονικής

Το πρότυπο Υπηρεσιοστρεφούς Αρχιτεκτονικής (*Service Oriented Architecture - SOA*) 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 διαδικτυακών υπηρεσιών οδηγεί στην υιοθέτηση της λεγόμενης Υπηρεσιοστρεφούς Αρχιτεκτονικής, οι ρόλοι και λειτουργίες της οποίας παρουσιάζονται στο παρακάτω σχήμα:



Σχήμα 11: Δομή Υπηρεσιοστρεφούς αρχιτεκτονικής

Η σελίδα αυτή είναι σκόπιμα λευκή

Η συγκεκριμένη αρχιτεκτονική υποδεικνύει μια σχέση εξυπηρετητή-πελάτη (server-client) ανάμεσα στον πάροχο υπηρεσιών («service provider»), που παίζει το ρόλο του εξυπηρετητή (server) και τον ζητώντα την υπηρεσία («service-requestor»), που παίζει το ρόλο του πελάτη (client). Ο πάροχος υπηρεσιών είναι αυτός που παρέχει την υπηρεσία δεχόμενος μηνύματα κλήσεις από τους πελάτες. Είναι επίσης υπεύθυνος για τη δημιουργία της περιγραφής της υπηρεσίας («service description») και τη δημοσίευσή της σε κάποιο κατάλογο-οδηγό υπηρεσιών (όπως για παράδειγμα ο Universal Description Discovery and Integration - UDDI). Ο πελάτης αναζητά την υπηρεσία και την περιγραφή της σε κάποιο κατάλογο υπηρεσιών («service registry») και στη συνέχεια καλεί κατάλληλα την επιθυμητή υπηρεσία.

4.2.4 Γλώσσα Περιγραφής Διαδικτυακών υπηρεσιών

Οι διαδικτυακές υπηρεσίες θα πρέπει να ορίζονται με συγκεκριμένο τρόπο έτσι ώστε να μπορούν να εντοπιστούν και χρησιμοποιηθούν από άλλες υπηρεσίες και εφαρμογές. Για αυτό το σκοπό δημιουργήθηκε από τον οργανισμό W3C («The World Wide Web Consortium») [50], μια γλώσσα περιγραφής γνωστή ως *Web Services Description Language* (WSDL) [47].

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 απαιτείται για την ακριβή χρήση μιας διαδικτυακής υπηρεσίας.

4.2.5 Πρωτόκολλο Επικοινωνίας

Το πρωτόκολλο Simple Object Access Protocol (SOAP) από το πρότυπο επικοινωνίας διαδικτυακών υπηρεσιών. Καθώς χρησιμοποιείται ευρέως παρατηρείται συχνά η παράφραση

του ακρωνύμιου από Simple Object Access Protocol σε Service-Oriented Architecture (or Application) Protocol.

Το συγκεκριμένο πρωτόκολλο διαμορφώνει ένα πρότυπο μήνυμα που αποτελείται από ένα XML έγγραφο ικανό να περιγράψει δεδομένα όπως RPC (Remote Procedure Call) κλήσεις. 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 στη SOA δομή, αφού με την βοήθεια της περιγραφής WSDL είναι εφικτή η επικοινωνία και συνεργασία οποιωνδήποτε υπηρεσιών στο δίκτυο.

Στο σχήμα της επόμενης σελίδας (Σχήμα 12) παρουσιάζεται η χρήση του πρωτοκόλλου σε εφαρμογές είτε για προτυποποιημένη RPC επικοινωνία είτε για γενική χρήση μεταφοράς μηνύματος.

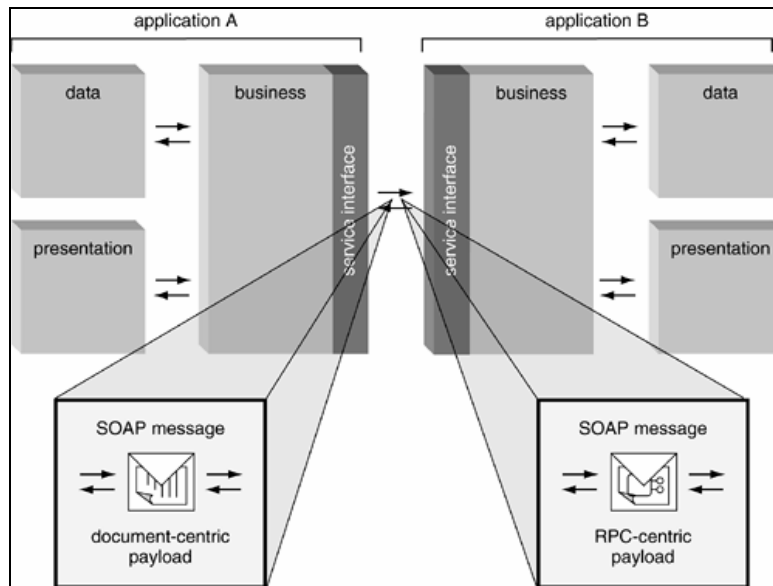
Το πρωτόκολλο SOAP καθορίζει ένα XML έγγραφο με μια συγκεκριμένη δομή που περιλαμβάνει όλες τις απαραίτητες πληροφορίες. Η βασική δομή ενός τέτοιου εγγράφου είναι η παρακάτω:

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

Εσωτερικά του “σώματος” (Body) του εγγράφου, ενσωματώνεται η πληροφορία που μεταφέρεται από τη μία υπηρεσία στην άλλη. Η τεχνολογία που χρησιμοποιείται για απομακρυσμένη κλήση υπηρεσιών είναι η SOAP-RPC (επέκταση της RPC) και η οποία εισάγει στο SOAP έγγραφο τις απαιτούμενες πληροφορίες, σύμφωνα και με το WSDL της υπηρεσίας που θα κληθεί και αποστέλλει το μήνυμα.

Επιπλέον, όπως φαίνεται και στο σχήμα της επόμενης σελίδας (Σχήμα 12), το πρωτόκολλο 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.



Σχήμα 12: Χρήση SOAP πρωτοκόλλου σε εφαρμογές

4.3 Αρχιτεκτονική Ανοιχτών Υπηρεσιών Πλέγματος

Οι συντονισμένες μελέτες του *Global Grid Forum* (GGF) [79] - ενός οργανισμού αποτελούμενου από χρήστες, προγραμματιστές, ερευνητές με σκοπό την ανάπτυξη προτύπων για την υποστήριξη των τεχνολογιών πλέγματος, το οποίο προέκυψε από τη συνένωση του *Open Grid Forum* (OGF) [76] με το επιχειρηματικό φόρουμ *Enterprise Grid Alliance* (EGA) [78] - κατέληξε το 2002 στην Αρχιτεκτονική Ανοιχτών Υπηρεσιών - *Open Grid Service Architecture* [35], 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.

4.3.1 Στόχοι της Αρχιτεκτονικής

Οι κύριοι στόχοι της αρχιτεκτονικής OGSA συνοψίζονται ακολούθως:

- Διαχείριση πόρων απομακρυσμένων ετερογενών συστημάτων

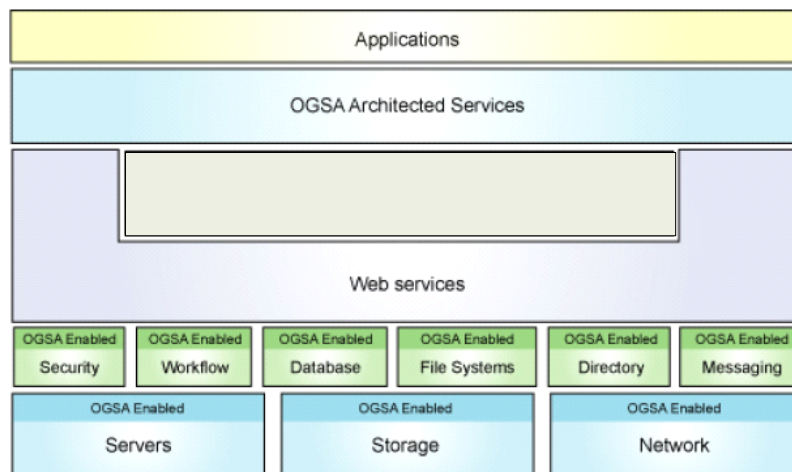
Η σελίδα αυτή είναι σκόπιμα λευκή

- Απόδοση σημαντικής ποιότητας στις υπηρεσίες που παρέχει (Quality of Service)
- Παροχή βάσεων για αυτόνομες λύσεις διαχείρισης. Η διαφορετικότητα των πόρων του πλέγματος και ο δυναμικός τρόπος συνεργασίας τους απαιτεί ένα σύστημα διαχείρισης που θα προσαρμόζεται σύμφωνα με τις εκάστοτε ανάγκες
- Καθορισμός γνωστών προτύπων και πρωτοκόλλων λειτουργίας. Η δυναμική εισαγωγή και συνεργασία ετερογενών συστημάτων σε περιβάλλον πλέγματος βασίζεται στην υιοθέτηση γνωστών, ευρείας χρήσης προτύπων
- Υιοθέτηση, εκμετάλλευση και προσαρμογή υπαρχουσών τεχνολογιών σε περιβάλλον πλέγματος (όπου αυτό είναι εφικτό) καθώς η αρχιτεκτονική OGSA βασίζεται στη τεχνολογία των διαδικτυακών υπηρεσιών (βλ. ενότητα 4.2.2).

4.3.2 Περιγραφή της Αρχιτεκτονικής

Η αρχιτεκτονική OGSA αποτελείται από τέσσερα βασικά επίπεδα (Σχήμα 13):

- Επίπεδο πόρων (φυσικών και λογικών)
- Επίπεδο υπηρεσιών διαδικτύου (Web Services)
- Επίπεδο υπηρεσιών αρχιτεκτονικά τοποθετημένων στην OGSA
- Επίπεδο εφαρμογών πλέγματος



Σχήμα 13: Αρχιτεκτονική OGSA

Η σελίδα αυτή είναι σκόπιμα λευκή

4.3.2.1 Επίπεδο φυσικών και λογικών πόρων

Αυτό είναι το πιο χαμηλό επίπεδο από τα τέσσερα και έχει να κάνει με τους πόρους που συμμετέχουν στο πλέγμα. Η έννοια του ‘πόρου’ στο OGSA είναι πολύ σημαντική και δεν είναι συγκεκριμένη. Ως πόρος (resource) σε ένα πλέγμα 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 δεδομένων, δικτυακές υποδομές κλπ. Αυτά που αναφέρθηκαν παραπάνω είναι κυρίως οι φυσικοί πόροι. Εκτός από αυτούς έχουμε και τους λογικούς, κάποια ενδιάμεσα συστήματα (middleware) που χρησιμοποιώντας τους φυσικούς πόρους, προσφέρουν στοιχειώδεις υπηρεσίες, όπως διαχείριση αρχείων ή βάσεων δεδομένων, στο παραπάνω επίπεδο.

4.3.2.2 Επίπεδο υπηρεσιών διαδικτύου (Web Services)

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

4.3.2.3 Επίπεδο υπηρεσιών πλέγματος

Οι υπηρεσίες διαδικτύου παρέχουν τη βασική υποδομή για το επόμενο επίπεδο, των υπηρεσιών πλέγματος που αναφέρονται σε υπηρεσίες υπολογισμού, πληροφοριών, πυρήνα,

κ.α. Για την υλοποίηση αυτών των υπηρεσιών περιγράφεται στην επόμενη ενότητα το Πλαίσιο Πόρων Διαδικτυακών Υπηρεσιών (WSRF). Βάσει των προδιαγραφών του προαναφερθέντος πλαισίου η αρχιτεκτονική OGSA παρουσιάζεται ως μια αρχιτεκτονική βασιζόμενη στις υπηρεσίες (πρόκειται δηλαδή για Υπηρεσιοστρεφή Αρχιτεκτονική - όπως περιγράφεται στην ενότητα 4.2).

4.3.2.4 Επίπεδο εφαρμογών πλέγματος

Αναφέρεται στο επίπεδο των εφαρμογών πλέγματος που χρησιμοποιούν μια ή περισσότερες υπηρεσίες πλέγματος του προηγούμενου επιπέδου.

4.3.2.5 Πλαίσιο Πόρων Διαδικτυακών Υπηρεσιών (WSRF)

Το Πλαίσιο Πόρων Διαδικτυακών Υπηρεσιών Ιστού (Web Service Resource Framework) [36] προτείνει το διαχωρισμό των διαδικτυακών υπηρεσιών με τις πληροφορίες κατάστασης. Η κατάσταση δεν τοποθετείται στη διαδικτυακή υπηρεσία (οπότε θα ονομαζόταν «stateful») αλλά διατηρείται σε μια χωριστή οντότητα αποκαλούμενη «πόρος», ο οποίος αποθηκεύει όλες τις πληροφορίες κατάστασης. Το πλαίσιο WSRF καθορίζει μια οικογένεια προδιαγραφών όπως περιγράφονται στη συνέχεια:

- WS-ResourceProperties: Η προδιαγραφή WS-ResourceProperties (Ιδιότητες Πόρων) καθορίζει πώς τα στοιχεία που συνδέονται με έναν stateful πόρο μπορούν να ρωτηθούν και να αλλάξουν χρησιμοποιώντας τις τεχνολογίες διαδικτυακών υπηρεσιών. Αυτό επιτρέπει την προσέγγιση από τους πελάτες με τυποποιημένα μέσα των στοιχείων εκείνων που θα συνδέονται με έναν WS-Resource. Η δήλωση των ιδιοτήτων των WS-Resources αντιπροσωπεύει μια προβολή ή μια άποψη σχετικά με την κατάσταση των WS-Resources. Αυτή η προβολή αντιπροσωπεύει έναν

υπονοούμενο τύπο πόρων που χρησιμεύει στο να καθορίσει μια βάση για πρόσβαση στις ιδιότητες των πόρων μέσω των διεπαφών των διαδικτυακών υπηρεσιών.

- WS-ResourceLifetime: Η προδιαγραφή WS-ResourceLifetime (Κύκλος Ζωής Πόρου) καθορίζει δύο τρόπους καταστροφής ενός WS-Resource: άμεσος και σχεδιασμένος. Αυτό επιτρέπει μια ευελιξία στους σχεδιαστές στο πώς οι εφαρμογές υπηρεσιών Ιστού τους μπορούν να καθαρίσουν από τους πόρους που δεν απαιτούνται πλέον.
- WS-BaseFaults: Η προδιαγραφή WS-BaseFaults (Βασικό Σφάλμα) καθορίζει έναν τύπο σχημάτων XML για βασικά σφάλματα, μαζί με τους κανόνες για το πώς αυτός ο τύπος σφαλμάτων χρησιμοποιείται από τις διαδικτυακές υπηρεσίες. Ένας σχεδιαστής μιας εφαρμογής διαδικτυακών υπηρεσιών χρησιμοποιεί συχνά διεπαφές που καθορίζονται από άλλους. Η διαχείριση των σφαλμάτων σε μια τέτοια εφαρμογή είναι δυσκολότερη όταν χρησιμοποιεί κάθε διεπαφή μια διαφορετική σύμβαση για την αντιπροσώπευση των κοινών πληροφοριών στα μηνύματα σφαλμάτων. Οποιαδήποτε υποστήριξη για τον προσδιορισμό και τη διευκρίνιση προβλημάτων στα μηνυμάτων σφαλμάτων των διαδικτυακών υπηρεσιών με έναν κοινό κι ενιαίο τρόπο μπορεί να ενισχύσει τη διαχείριση σφαλμάτων. Όταν οι διαθέσιμες πληροφορίες σφαλμάτων από τις διάφορες διεπαφές είναι πλήρεις, είναι ευκολότερο για τους αιτούντες να καταλάβουν το είδος των σφαλμάτων.
- WS-ServiceGroup: Η προδιαγραφή WS-ServiceGroup (Ομάδα Υπηρεσιών) καθορίζει τα μέσα με τα οποία οι διαδικτυακές υπηρεσίες και οι WS-Resources μπορούν να αθροιστούν ή να συγκεντρωθούν για έναν συγκεκριμένο σκοπό. Η συμμετοχή στην ομάδα πρέπει να περιοριστεί με κάποιο τρόπο προκειμένου οι αιτούντες να διαμορφώνουν ερωτήσεις με σημασία προς το περιεχόμενο του συνόλου των υπηρεσιών. Οι περιορισμοί συμμετοχής μέλους εκφράζονται με βάση την πρόθεση, χρησιμοποιώντας έναν μηχανισμό ταξινόμησης. Περαιτέρω, τα μέλη κάθε πρόθεσης

πρέπει να μοιραστούν ένα κοινό σύνολο πληροφοριών πάνω στο οποίο οι ερωτήσεις μπορούν να εκφραστούν.

- WS-Notification: Η υπηρεσία αυτή (δεν περιέχεται στις προδιαγραφές) επιτρέπει σε μια διαδικτυακή υπηρεσία να διαμορφωθεί ως παραγωγός ειδοποιήσεων, ενόσω ορισμένοι πελάτες να είναι οι καταναλωτές (ή συνδρομητές) αυτών των ειδοποιήσεων. Αυτό σημαίνει ότι εάν μια αλλαγή εμφανίζεται στη διαδικτυακή υπηρεσία (ή, πιο συγκεκριμένα, σε έναν από τους WS-Resources), η αλλαγή αυτή θα κοινοποιηθεί σε όλους τους συνδρομητές. (Δεν κοινοποιούνται όλες οι αλλαγές, παρά μόνο εκείνες που επιλέγονται από τον προγραμματιστή διαδικτυακών υπηρεσιών).
- WS-Addressing: Η υπηρεσία αυτή (επίσης δεν περιέχεται στις προδιαγραφές) παρέχει έναν μηχανισμό για διευθυνσιοδότηση των διαδικτυακών υπηρεσιών που είναι πιο ευπροσάρμοστος από τα απλά παγκόσμια αναγνωριστικά (Universal Resource Identifier, URI). Με τον τρόπο αυτό είναι δυνατή η χρήση της WS-Addressing για τη διευθυνσιοδότηση ενός ζεύγους διαδικτυακής υπηρεσίας και πόρου (ένας WS-Resource).

4.4 Καινοτόμα Αρχιτεκτονική Πλέγματος Επόμενης Γενιάς

Στη συγκεκριμένη ενότητα αναφέρονται στοιχεία για την αρχιτεκτονική σχεδίαση του πλέγματος επόμενης γενιάς (Next Generation Grid Architecture) όπως αυτή διαμορφώνεται από το ερευνητικό έργο NextGRID [65] και συγκεκριμένα στοιχεία της αρχιτεκτονικής που επιτρέπουν το σχεδιασμό, περιγραφή και εκτέλεση δυναμικών ροών εργασίας.

Το NextGRID (Architecture for Next Generation Grids) είναι ένα ερευνητικό έργο στο οποίο συμμετέχουν ευρωπαϊκά ερευνητικά και ακαδημαϊκά ιδρύματα καθώς και επιχειρήσεις που δραστηριοποιούνται στο χώρο της τεχνολογίας πλέγματος. Στόχος του έργου είναι η

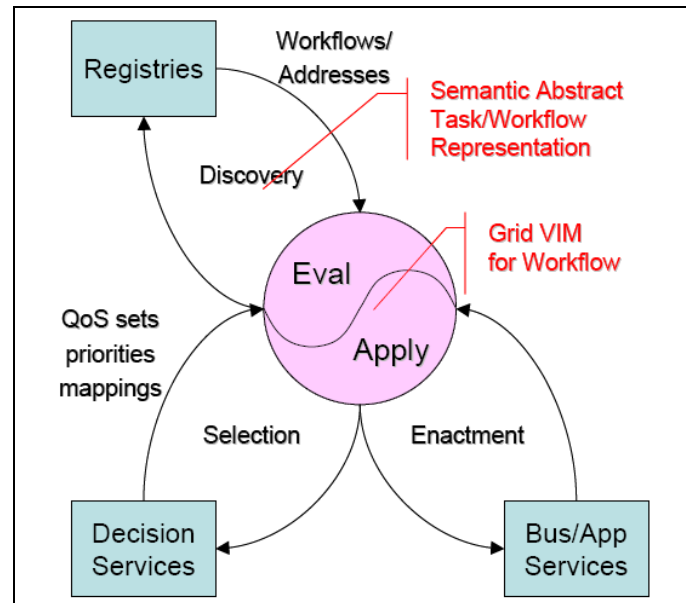
ανάλυση και ο σχεδιασμός της αρχιτεκτονικής του πλέγματος επόμενης γενιάς [93], [94], [95]. Μεταξύ των στόχων του NextGRID περιλαμβάνονται η εξάπλωση της χρήσης της τεχνολογίας πλέγματος από τον ακαδημαϊκό και ερευνητικό χώρο σε εφαρμογές του επιχειρηματικού κόσμου με έμφαση στην ασφάλεια και το δυναμικό χαρακτήρα των στοιχείων αυτού.

Στα πλαίσια του έργου γίνεται μελέτη, σχεδιασμός και ανάλυση μιας σειράς υπηρεσιών πλέγματος που θα προσφέρουν λειτουργικότητα σε διάφορους τομείς όπως την εύρεση υπηρεσιών (service discovery), τα μητρώα υπηρεσιών (registries), την περιγραφή υπηρεσιών με χρήση σημασιολογικής (semantic) γλώσσας, τα συμβόλαια παροχής υπηρεσιών SLAs (βλ. ενότητα 2.4.6) κ.α. Επιπλέον το NextGRID εισάγει ένα επιχειρηματικό μοντέλο στο οποίο κεντρικό ρόλο έχουν οι ροές εργασίας, το οποίο περιγράφεται αναλυτικά στη συνέχεια της παρούσας διατριβής

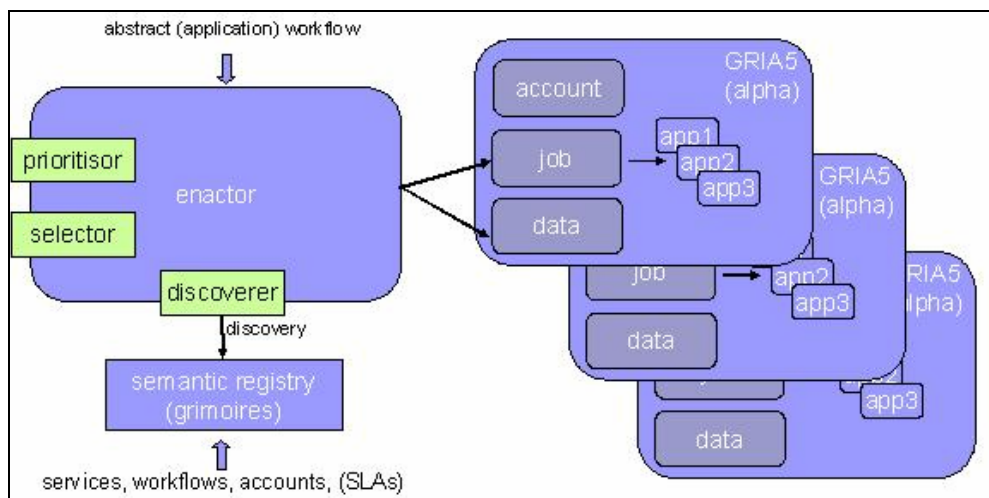
4.4.1 Μοντέλο Εικονικής Υποδομής

Σε ένα περιβάλλον πλέγματος, οι πάροχοι των υπηρεσιών μπορούν να προβάλλουν λεπτομερείς ροές εργασίας που περιγράφουν τις αλληλεπιδράσεις και τις προϋποθέσεις που είναι απαραίτητες να πληροί ένας πελάτης για να χρησιμοποιήσει τις υπηρεσίες τους. Υπάρχει όμως η πιθανότητα, δύο πάροχοι υπηρεσιών να μη χρησιμοποιούν την ίδια υποδομή πλέγματος και να έχουν πολύ διαφορετικές επιχειρησιακές (business) πολιτικές. Αυτό μπορεί να ισχύει ακόμη και αν προσφέρουν τις ίδιες υπηρεσίες και λειτουργίες στους πιθανούς πελάτες. Έτσι προκύπτει πρόβλημα για τους σχεδιαστές εφαρμογών και ροών εργασίας, όταν επιθυμούν η εφαρμογή τους να επιτύχει ένα λειτουργικό στόχο χρησιμοποιώντας στο χρόνο εκτέλεσης μία από τις δύο υπηρεσίες που προσφέρουν οι δύο πάροχοι των υπηρεσιών. Πρόκειται για ένα μοντέλο που στοχεύει στην αφαίρεση τόσο των επιχειρησιακών

διεργασιών, όσο και των απαραίτητων αλληλεπιδράσεων με την υποδομή πλέγματος με τη χρήση προσαρμοσμένων ροών εργασίας (adaptive workflows) και τη δυναμική δέσμευσης υπηρεσιών που περιλαμβάνει σημασιολογική εύρεση υπηρεσιών (service discovery) και επιλογή υπηρεσιών (service selection). Η λογική της εφαρμογής μπορεί να αναπαρασταθεί με αφηρημένες ροές εργασίας, οι οποίες περιλαμβάνουν λειτουργικούς περιορισμούς της εφαρμογής. Κατά τη διάρκεια της εκτέλεσης μιας ροής, αφηρημένες εργασίες της εφαρμογής μετατρέπονται σε συγκεκριμένες (για τις οποίες είναι γνωστή η υλοποίησή τους), μέσω μιας διαδικασίας που περιλαμβάνει εύρεση (discovery) και επιλογή (selection) υπηρεσιών πριν την εκτέλεση της εργασίας. Το Μοντέλο Εικονικής Υποδομής (Grid Virtual Infrastructure Model - GridVIM) και τα κύρια αρχιτεκτονικά μέρη του φαίνονται στο σχήμα της επόμενης σελίδας (Σχήμα 14). Βάσει όσων παρουσιάζονται στο σχήμα, η λειτουργία του μοντέλου χωρίζεται σε δύο κύκλους, την εκτίμηση (evaluate) και την εφαρμογή (apply). Η φάση εκτίμησης συνδέει μια περιγραφή αφηρημένης υπηρεσίας (abstract service description) σε μία υλοποίηση διαδικτυακής υπηρεσίας ή σε μια νέα αφηρημένη ροή εργασίας. Αρχικά γίνεται εύρεση (discovery) των υποψήφιων υπηρεσιών και ροών εργασίας πραγματοποιώντας ερωτήματα προς διάφορα μητρώα (registries). Στη συνέχεια γίνεται χρήση των υπηρεσιών αποφάσεων (decision services) που καθορίζουν, συνδυάζοντας οργανωτικές πολιτικές και την ποιότητα υπηρεσίας, ποια από τις υποψήφιες υπηρεσίες θα επιλεγεί. Η φάση εκτίμησης που ακολουθεί, εκτελεί πλέον τις συγκεκριμένες διεργασίες (concrete processes) που πλέον συνοδεύονται από τις πληροφορίες εκτέλεσής τους, όπως φαίνεται στο Σχήμα 15.



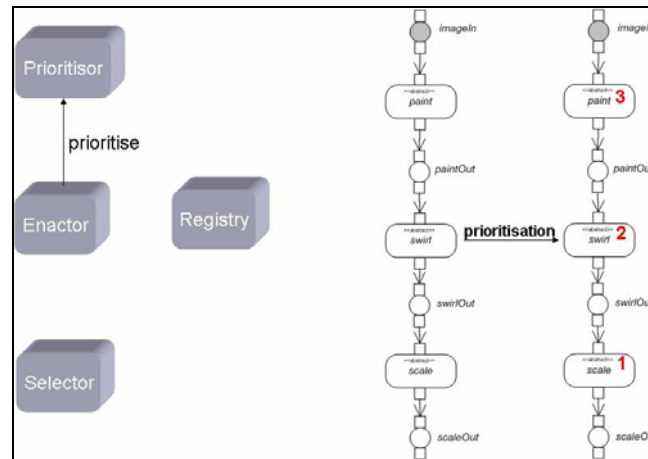
Σχήμα 14: Αρχιτεκτονική του μοντέλου GridVIM



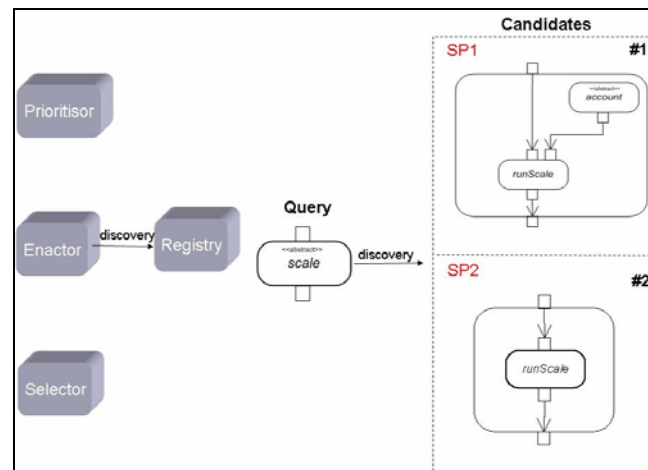
Σχήμα 15: Μοντέλο GridVIM

Επιπρόσθετα στα σχήματα που ακολουθούν (Σχήματα 16, 17 και 18) παρουσιάζονται γραφικά οι τρεις (3) φάσεις καθορισμού προτεραιοτήτων (prioritisor), εύρεσης (discoverer) και επιλογής (selector) υπηρεσιών πλέγματος.

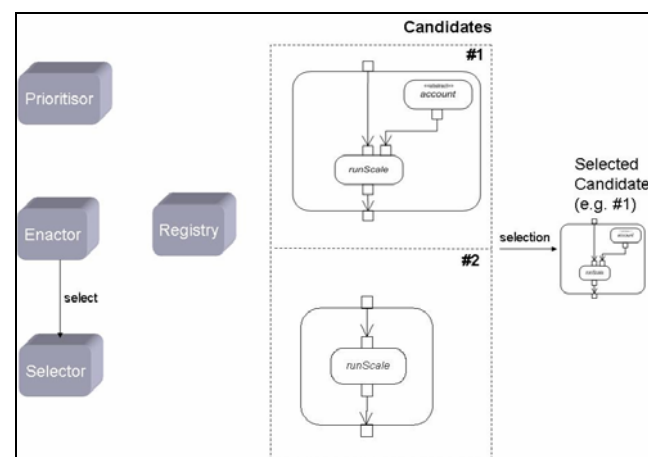
Η σελίδα αυτή είναι σκόπιμα λευκή



Σχήμα 16: Καθορισμός προτεραιοτήτων υπηρεσιών



Σχήμα 17: Εύρεση υπηρεσιών



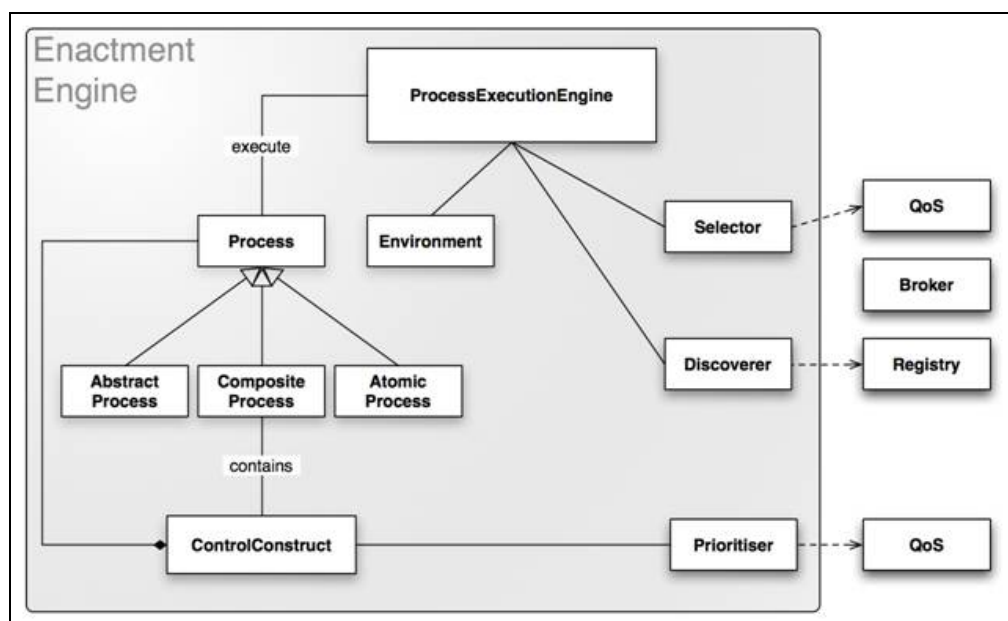
Σχήμα 18: Επιλογή υπηρεσιών

Η σελίδα αυτή είναι σκόπιμα λευκή

Η παραπάνω προσέγγιση επιτρέπει σε εφαρμογές να προσαρμόζονται ώστε να μπορούν να λειτουργήσουν σε διαφορετικά περιβάλλοντα πλέγματος, τόσο από πλευράς υποδομής όσο και από πλευράς επιχειρησιακών διεργασιών (business process). Έτσι κάθε φορά που μια εφαρμογή πλέγματος εκτελείται, μπορεί να φορτώνει μια συνολική πολιτική ρυθμίσεων του μοντέλου GridVIM, η οποία παρέχει πληροφορίες για το πώς η εφαρμογή θα προσαρμοστεί για να εκτελεστεί στο όποιο επιχειρησιακό περιβάλλον.

4.4.2 Συγκεκριμένα Στοιχεία Αρχιτεκτονικής

Αρχιτεκτονικά το μοντέλο GridVIM μπορεί να χωριστεί (όπως φαίνεται στο παρακάτω σχήμα - Σχήμα 19) στον GridVIM Enactor και σε υπηρεσίες πλέγματος με τις οποίες αλληλεπιδρά όπως η υπηρεσία διασφάλισης ποιότητας (QoS) με χρήση μηχανισμών απεικόνισης ροών εργασίας, η υπηρεσία διαμεσολάβησης (Broker) και η υπηρεσίας καταλόγου (Registry). Η επικοινωνία με αυτές τις υπηρεσίες γίνεται μέσω των αντίστοιχων διεπαφών Selector, Discoverer και Prioritiser των οποίων η λειτουργία θα αναλυθεί στις επόμενες ενότητες.



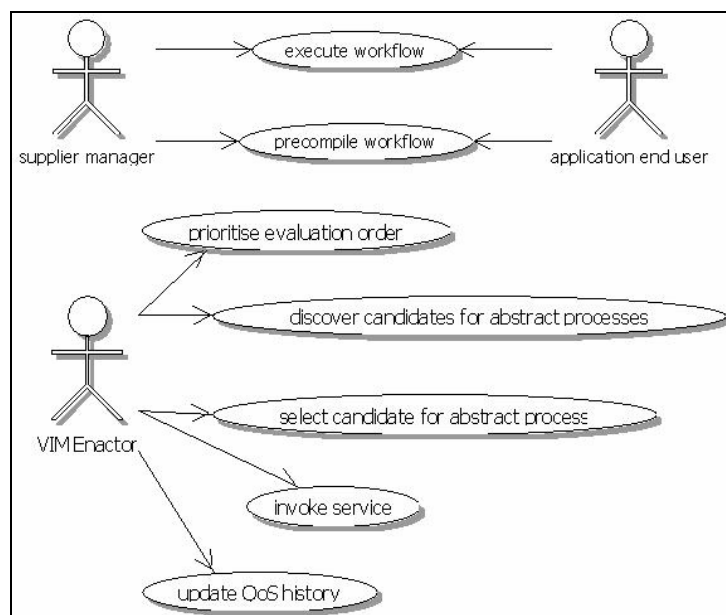
Σχήμα 19: Δομικά στοιχεία του GridVIM

Η σελίδα αυτή είναι σκόπιμα λευκή

4.4.2.1 GridVIM Enactor

Ο GridVIM enactor είναι ένας μηχανισμός εκτέλεσης προσαρμοστικών ροών εργασίας (εκτελεί σημασιολογικές ροές που περιγράφονται με τη γλώσσα OWL-WS). Ο δυναμικός χαρακτήρας του μηχανισμού και των ροών εργασίας επιτυγχάνεται με την εκτίμηση της ροής κατά το χρόνο εκτέλεσης, η οποία στοχεύει στο να το προσαρμόσει σύμφωνα με τις επιχειρησιακές διεργασίες, την πολιτική εικονικών οργανισμών και το περιβάλλον. Όπως αναφέρθηκε σε προηγούμενη ενότητα η εκτέλεση των ροών εργασίας βασίζεται σε κύκλους εκτίμησης-εφαρμογής (βλ. ενότητα 4.4.1).

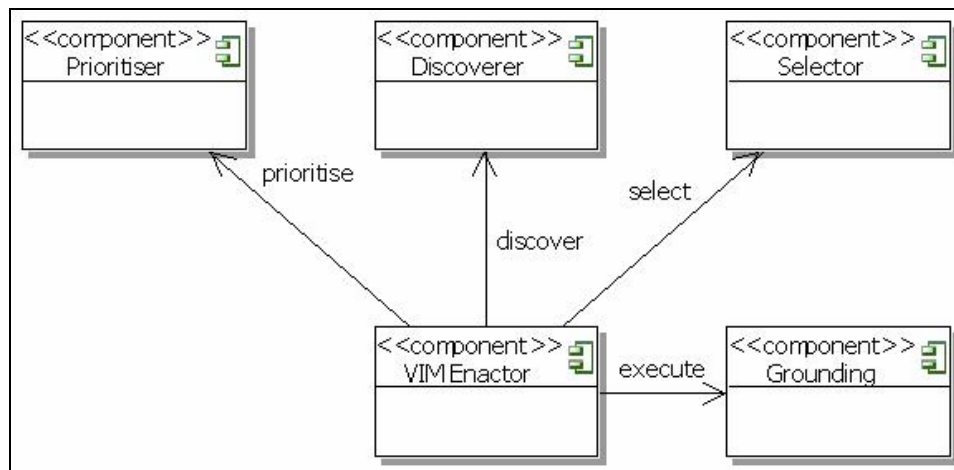
Η προαναφερθείσα εκτίμηση αρχικοποιείται όποτε μια αφηρημένη διεργασία (Abstract Process) πρέπει να εκτιμηθεί. Η διαδικασία της εκτίμησης αποτελείται από τη εύρεση (discovery), κατά την οποία αναζητούνται υποψήφιες υλοποιήσεις της αφηρημένης διεργασίας και από την επιλογή (selection), κατά την οποία επιλέγεται μία από τις υλοποιήσεις. Στη συνέχεια περιγράφονται αναλυτικά οι προαναφερθείσες υπηρεσίες και οι διεπαφές με το GridVIM enactor. Στη συνέχεια περιγράφονται σχηματικά τα σενάρια χρήσης του GridVIM enactor (Σχήμα 20) με τις βασικές λειτουργίες που δύναται να ολοκληρώσει:



Σχήμα 20: Περιπτώσεις χρήσης

Η σελίδα αυτή είναι σκόπιμα λευκή

Οι παραπάνω λειτουργίες όπως εκφράζονται από τις περιπτώσεις χρήσης υλοποιούνται με τη χρήση των διεπαφών και των υπηρεσιών που περιγράφηκαν αναλυτικά στην προηγούμενη ενότητα και τα οποία παρουσιάζονται στο παρακάτω σχήμα – Component Diagram (Σχήμα 21):



Σχήμα 21: Βασικά στοιχεία και λειτουργίες δομής GridVIM

4.4.2.2 Η υπηρεσία και η διεπαφή καθορισμού προτεραιοτήτων (Prioritiser)

Η υπηρεσία αυτή χρειάζεται για να καθοριστεί η σειρά με την οποία γίνεται η εκτέλεση της ροής εργασίας, η σειρά δηλαδή με την οποία θα γίνει προσπάθεια οι αφηρημένες διεργασίες να μετατραπούν σε συγκεκριμένες. Μόνο οι αφηρημένες διεργασίες ενός workflow θα εκτιμηθούν, οπότε η χρήση της υπηρεσίας αυτής γίνεται μόνο για τις διεργασίες αυτές. Όλα στιγμιότυπα των διεργασιών σε OWL-WS έχουν μια παράμετρο που ονομάζεται Priority (προτεραιότητα). Η παράμετρος αυτή έχει ως τιμή έναν ακέραιο αριθμό, ο οποίος καθορίζει την προτεραιότητα της διεργασίας για την εκτίμηση. Ειδικότερα, όσο μεγαλύτερη είναι η τιμή της παραμέτρου, τόσο πιο χαμηλή είναι η προτεραιότητα της διεργασίας. Για να τεθούν οι προτεραιότητες σε μία δομή ελέγχου, η υλοποίηση της Prioritiser θα πρέπει να εξετάζει τις διεργασίες μιας δομής ελέγχου ροής εργασίας και να θέτει μια τιμή προτεραιότητας σε εκείνες τις διεργασίες που είναι στιγμιότυπα της διεπαφής. Οι εντολές με τις οποίες τίθεται η

Η σελίδα αυτή είναι σκόπιμα λευκή

προτεραιότητα σε μια αφηρημένη διεργασία είναι:

```
Priority priority=ont.createPriority(priorityURI);

//Καθορισμός της τιμής της ιδιότητας

OWLDatatype dataType = ont.getDataType(new
URI("http://www.w3.org/2001/XMLSchema#integer"));

priority.setParamType(dataType);

EntityFactory ef=new EntityFactory();

OWLDatatype priorityValue = ef.createDataValue(1);

priority.setValue(priorityValue);

//Ανάθεση της προτεραιότητας στη διεργασία

priority.setProcess(process);

process.setPriority(priority);

Η διεπαφή διαθέτει μόνο μια μέθοδο, που δέχεται ως είσοδο μια λίστα δομών ελέγχου και δίνει ως
έξοδο τη λίστα αυτή, έχοντας όμως θέσει προτεραιότητες:

public interface Prioritiser{

public ControlConstructList prioritise(ControlConstructList ccList);

}
```

4.4.2.3 Η υπηρεσία και η διεπαφή εύρεσης υπηρεσιών (Discovery)

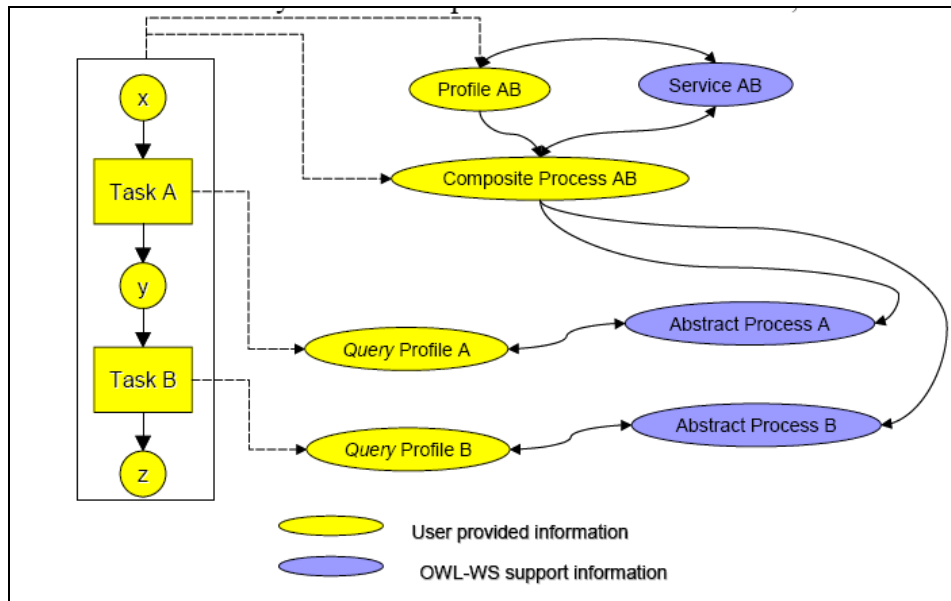
Η υπηρεσία Discoverer αναλαμβάνει τη λειτουργία της εύρεσης αποστέλλοντας για κάθε αφηρημένη διεργασία το προφίλ ερώτησής της (Query Profile) σε ένα μητρώο (Registry). Οι διεργασίες που παρέχονται ως αποτέλεσμα μπορούν να είναι αφηρημένες, συγκεκριμένες, ατομικές ή σύνθετες (δηλαδή ροές εργασίας). Το μητρώο υπηρεσιών (Service Registry) θα πρέπει να πραγματοποιεί ερωτήσεις στη βάση δεδομένων του, ώστε να βρίσκει υπηρεσίες ή

ροές εργασίας, των οποίων η λειτουργικότητα ταιριάζει στο προφίλ της αφηρημένης διεργασίας.

Η διεπαφή του μηχανισμού εύρεσης αποτελείται από μια μέθοδο εύρεσης, η οποία δέχεται ως είσοδο ένα προφίλ ερώτησης και επιστρέφει ως έξοδο μια λίστα από αντικείμενα διεργασιών, που μπορούν να υλοποιήσουν την αφηρημένη διεργασία. Επίσης είναι δυνατόν να υπάρχει και δεύτερη μέθοδος εύρεσης, η οποία δέχεται ως είσοδο μια αφηρημένη διεργασία και μπορεί να χρησιμοποιήσει είτε το προφίλ ερώτησής της, είτε άλλες πληροφορίες της, ώστε να πραγματοποιήσει σχετική ερώτηση προς το μητρώο. Η διεπαφή υλοποιείται ως εξής:

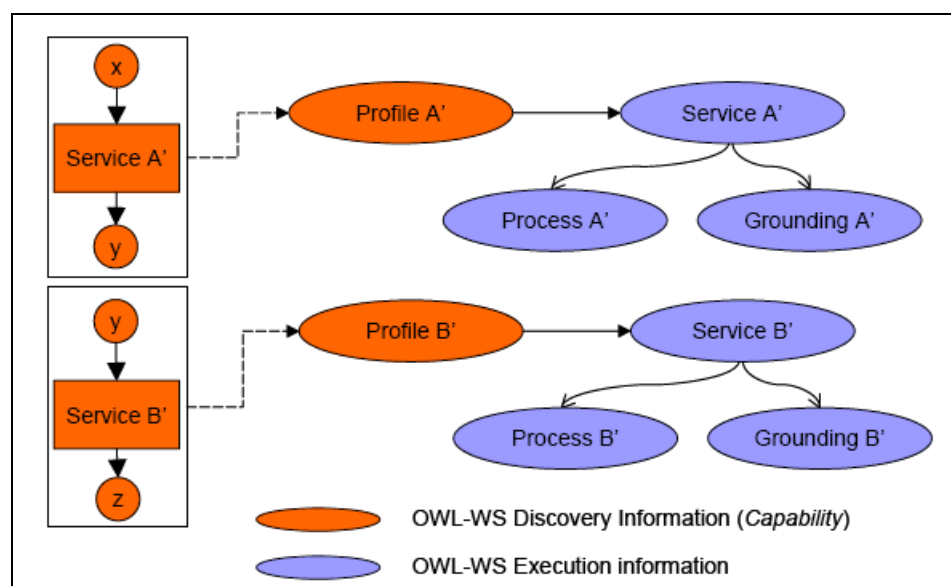
```
public interface ProcessDiscoverer {  
  
public List<org.mindswap.owls.process.Process> discover(Profile profile) throws DiscoveryException;  
  
public List<org.mindswap.owls.process.Process> discover(AbstractProcess process) throws  
DiscoveryException;  
  
}
```

Για παράδειγμα αναφέρεται η περίπτωση μιας αφηρημένης ροής εργασίας, που αποτελείται από δύο αφηρημένες διεργασίες A και B και οι οποίες συνοδεύονται από τα αντίστοιχα προφίλ ερώτησης (τα οποία περιέχουν λειτουργικές και μη απαιτήσεις για κάθε εργασία) με βάση τις οποίες θα γίνει αναζήτηση των πιθανών υποψήφιων. Τα παραπάνω παρουσιάζονται γραφικά στο σχήμα της επόμενης σελίδας (Σχήμα 22).



Σχήμα 22: Ερωτήσεις προφίλ

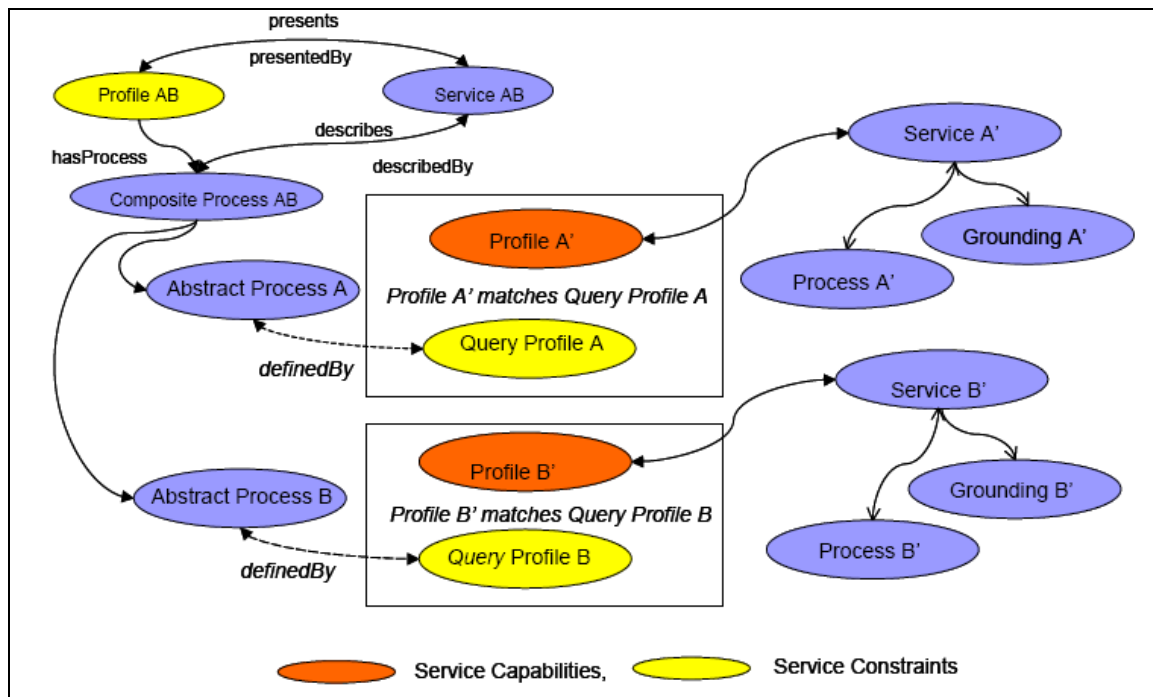
Στο επόμενο σχήμα (Σχήμα 23) παρουσιάζεται το προφίλ που περιλαμβάνει τις δυνατότητες της προσφερόμενης υπηρεσίας, ενώ η θεμελίωση αυτής περιέχει πληροφορίες που αφορούν την εκτέλεση της. Ουσιαστικά η εύρεση επιχειρεί να ταιριάζει τα προφίλ ερώτησης με τα προφίλ δυνατοτήτων των παρεχόμενων υπηρεσιών, έτσι ώστε να βρεθούν οι κατάλληλες διεργασίες που να υλοποιούν την αφηρημένη διεργασία για την οποία γίνεται η εύρεση. Το τελευταίο φαίνεται καλύτερα στο Σχήμα 24.



Σχήμα 23: Δυνατότητες προφίλ

Η σελίδα αυτή είναι σκόπιμα λευκή

Στο σχήμα που ακολουθεί (Σχήμα 24) φαίνεται η αντιστοίχιση του προφίλ ερώτησης με ένα προφίλ υπηρεσίας, αλλά γενικά τα προφίλ πολλών (υποψήφιων) υπηρεσιών μπορεί να αντιστοιχηθεί στο προφίλ μιας ερώτησης, οπότε στη συνέχεια θα πρέπει να γίνεται επιλογή μετά μιας από τις υποψήφιες υπηρεσίες.



Σχήμα 24: Αντιστοίχιση προφίλ στις ερωτήσεις

4.4.2.4 Η διεπαφή επιλογής (Selection)

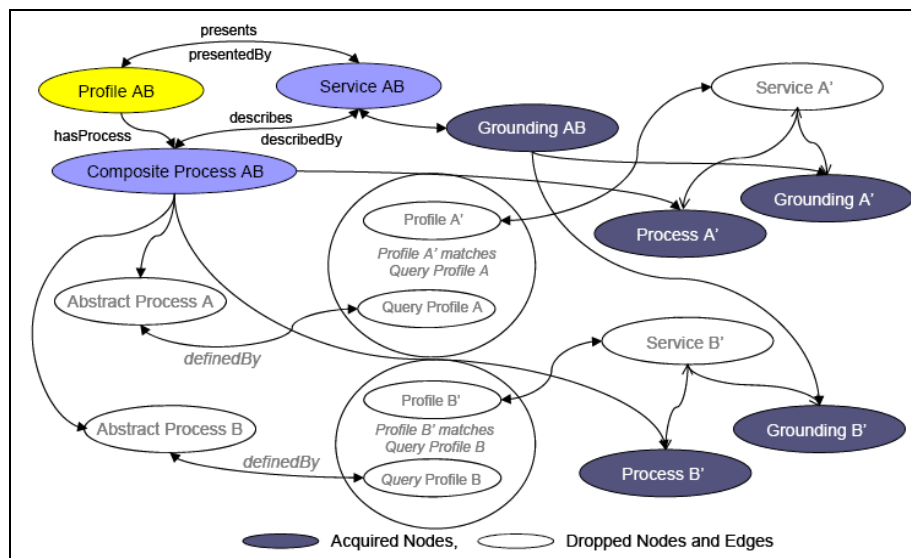
Η επιλογή (selection) είναι η διαδικασία, που καθορίζει τη διεργασία, η οποία θα υλοποιήσει την αφηρημένη διεργασία, για την οποία γίνεται εκτίμηση. Χρησιμοποιεί για αυτό το λόγο τις υπηρεσίες Broker και αντιστοίχισης ροών εργασίας που περιγράφεται αναλυτικά στο Κεφάλαιο 6 της διατριβής.

Κάθε αφηρημένη διεργασία διαθέτει μια ιδιότητα, η οποία παίρνει ως τιμή μια διεργασία, που λειτουργεί ως δείκτης προς τη διεργασία, που υλοποιεί την αντίστοιχη αφηρημένη. Επομένως, η διεπαφή αναλαμβάνει να θέσει την ιδιότητα αυτή για κάθε αφηρημένη διεργασία. Αυτό μπορεί να γίνει με δύο τρόπους:

Η σελίδα αυτή είναι σκόπιμα λευκή

- Επιλογή της υποψήφιας διεργασίας για κάθε αφηρημένη διεργασία χωριστά. Σε αυτή την περίπτωση είσοδος είναι η αφηρημένη διεργασία, στην οποία αφού τεθεί στην ιδιότητα επιστρέφεται ως έξοδος. Υπάρχουν δύο μέθοδοι που πραγματοποιούν τη λειτουργία αυτή. Η μία από τις δύο παίρνει ως επιπλέον είσοδο μια παράμετρο, που δείχνει εάν κάποια από τις υποψήφιες διεργασίες έχει παρουσιάσει ήδη πρόβλημα, έτσι ώστε να μην ξαναεπιλεγεί.
- Επιλογή των υποψήφιων διεργασιών για όλες τις αφηρημένες διεργασίες. Η μέθοδος αυτή επιχειρεί με μία κλήση της, να πραγματοποιήσει την επιλογή για όλες τις αφηρημένες διεργασίες μιας ροής εργασίας. Παίρνει ως είσοδο ένα αντικείμενο και για κάθε αφηρημένη διεργασία, που βρίσκει σε αυτό, πραγματοποιεί επιλογή.

Στο παρακάτω σχήμα (Σχήμα 25) παρουσιάζονται τα αποτελέσματα της επιλογής σε μια ροή εργασίας. Οι αφηρημένες διεργασίες αντικαθίστανται από συγκεκριμένες διεργασίες, που συνοδεύονται από την αντίστοιχη θεμελίωση (Grounding). Επίσης δημιουργείται και μια θεμελίωση (Grounding AB) για το σύνολο της, συνδυάζοντας στοιχεία από τις επιμέρους θεμελιώσεις, ενώ τα στοιχεία που δε χρειάζονται πλέον απορρίπτονται.



Σχήμα 25: Αποτελέσματα αντιστοίχισης

Η σελίδα αυτή είναι σκόπιμα λευκή

Επιπρόσθετα, είναι δυνατόν η επιλογή, να χρησιμοποιεί το μηχανισμό των προτεραιοτήτων, έτσι ώστε να καθορίσει τη σειρά επιλογής. Η διεπαφή έχει ως εξής:

```
public interface ProcessSelector {

    public AbstractProcess select(AbstractProcess process, org.mindswap.owls.process.Process exception)
    throws SelectionException;

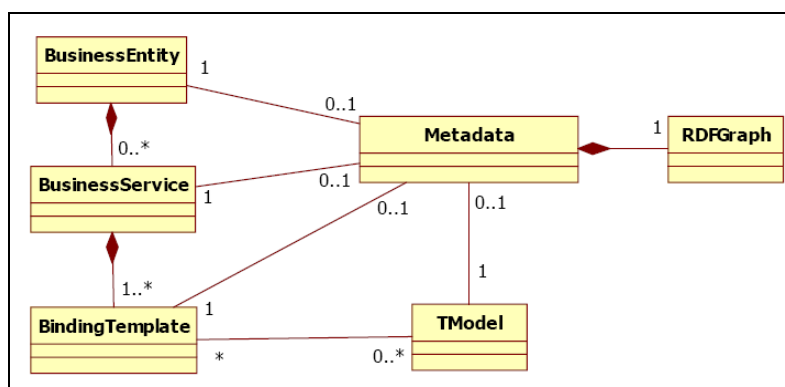
    public AbstractProcess select(AbstractProcess process) throws SelectionException;

    public ControlConstructList select(ControlConstructList ccList) throws SelectionException;

}
```

4.4.2.5 Η υπηρεσία Μητρώου

Η υπηρεσία μητρώου (Registry Service) του μοντέλου GridVIM απαιτείται για την υποστήριξη της δημοσίευσης περιγραφών ροών εργασίας, υπηρεσιών και εφαρμογών πλέγματος καθώς επίσης και για την υποστήριξη πραγματοποίησης ερωτημάτων προς τα πολύπλοκα, ημιδομημένα δεδομένα, που περιγράφουν τις οντότητες αυτές και περιέχουν πληροφορίες για τις λειτουργικές και μη λειτουργικές ιδιότητές τους, συμπεριλαμβανομένων και των πληροφοριών που σχετίζονται με τα συμβόλαια παροχής υπηρεσιών (βλ. ενότητα 2.4.6). Το μοντέλο GridVIM χρησιμοποιεί ένα μητρώο τύπου Grimoires [51], οι οντότητες του οποίου παρουσιάζεται στο παρακάτω διάγραμμα (Σχήμα 26):

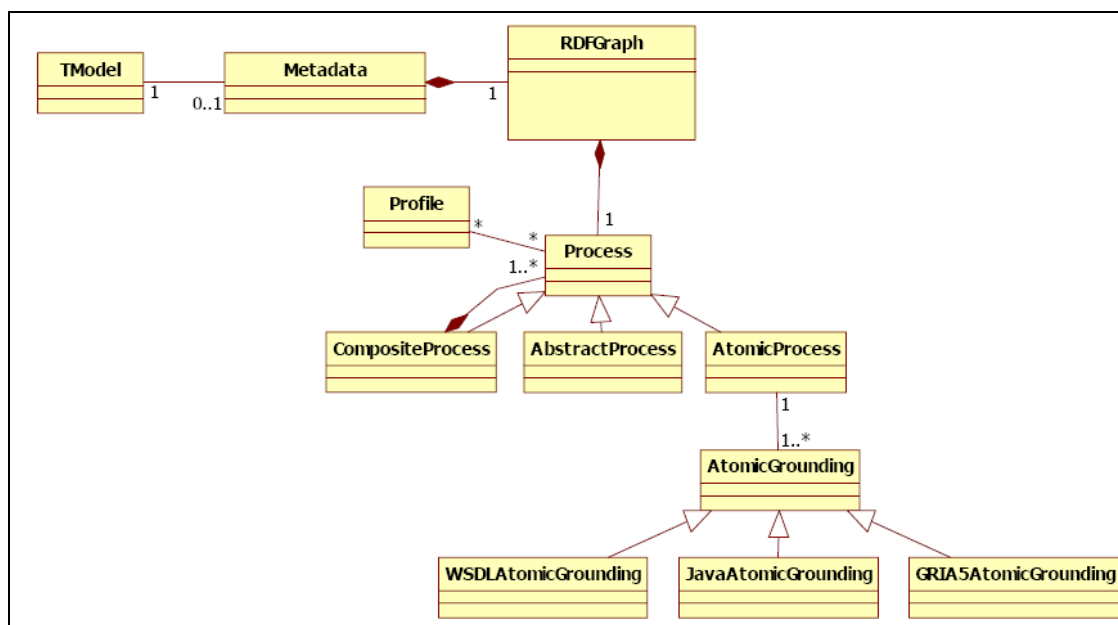


Σχήμα 26: Υπηρεσία μητρώου

Η σελίδα αυτή είναι σκόπιμα λευκή

Το συγκεκριμένο μητρώο υποστηρίζει σημασιολογικά ερωτήματα κάνοντας τη χρήση της γλώσσας RDQL [52], προσφέρει πληθώρα ρυθμιστικών ενεργειών όσον αφορά τον τρόπο διατήρησης των δεδομένων, συμπεριλαμβάνοντας και υποστήριξη για σχεσιακές βάσεις δεδομένων. Υποστηρίζει επίσης σχήματα εξουσιοδότησης-πιστοποιητικών, μηνύματα SOAP (βλ. ενότητα 4.2.5), συμφωνία με τα πρότυπα του WS-Security (βλ. ενότητα 4.3.2.5) και την ύπαρξη μεταδεδομένων για τις δημοσιευμένες οντότητες.

Επιπλέον απαιτείται προσαρμογή του μητρώου στο μοντέλο της γλώσσας OWL-WS για περιγραφή υπηρεσιών και ροών εργασίας. Η προσαρμογή που επιλέχθηκε ήταν κάθε οντότητα TModel να αντιστοιχιστεί με ένα διαφορετικό συνδυασμό οντοτήτων διεργασίας (Process), υπηρεσίας (Service), προφίλ (profile) και θεμελίωσης (Grounding) της OWL-WS. Κάθε τέτοιος συνδυασμός συνοδεύεται και από μία οντότητα μεταδεδομένων (Metadata), που μπορεί να περιέχει πληροφορίες για το συγκεκριμένο συνδυασμό οντοτήτων.



Σχήμα 27: Περιγραφή μοντέλου

Με τον τρόπο αυτό, χρειάζεται μόνο μια κλήση προς το μητρώο για να ληφθεί μια διεργασία μαζί με το προφίλ της. Η δημοσίευση μιας νέας υπηρεσίας γίνεται με δύο κλήσεις. Μία για να δημιουργηθεί μία οντότητα TModel και μία για να σχετιστεί με μια οντότητα Metadata.

Η σελίδα αυτή είναι σκόπιμα λευκή

Διαπιστώνεται εύκολα ότι για κάθε διαφορετική διεργασία, ακόμη και αν ανήκει στην ίδια υπηρεσία με μια άλλη, δημιουργείται ένα προφίλ, γεγονός που σημαίνει ότι είναι δυνατόν ένα προφίλ να παρουσιάζεται πολλές φορές, κάτι που εισάγει πλεονασμό στο μοντέλο.

Τέλος, για την εισαγωγή του μοντέλου αυτού στο GridVIM χρειάστηκε ο σχεδιασμός μιας διεπαφής για τη διαχείριση του μητρώου. Η διεπαφή αυτή διαθέτει τις παρακάτω μεθόδους:

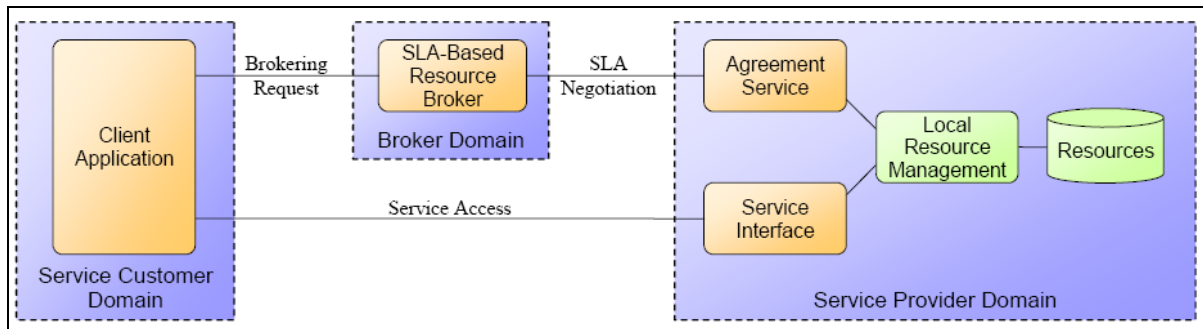
- *register*. Αυτή η μέθοδος δέχεται ως είσοδο μια διεργασία, η οποία θα καταχωρηθεί στο μητρώο. Η διεργασία συνοδεύεται προφανώς και από στοιχεία όπως το προφίλ της και η θεμελίωσή της που επίσης αποθηκεύονται στο μητρώο.
- *unregister*. Αυτή η μέθοδος εκτελεί την αντίθετη λειτουργία από αυτή της προηγούμενης μεθόδου.
- *getProcess*. Αυτή η μέθοδος δέχεται ως είσοδο το αναγνωριστικό μιας διεργασίας και ως έξοδο δίνει τη διεργασία αυτή.
- *getQueryCapabilities*. Η κλήση αυτής της μεθόδου έχει ως αποτέλεσμα την επιστροφή ενός αντικειμένου με πληροφορίες για τα είδη των ερωτημάτων που υποστηρίζει το μοντέλο.
- *sparqlSelectQuery* και *sparqlConstructQuery*. Οι μέθοδοι αυτές χρησιμοποιούν τη γλώσσα ερωτημάτων SPARQL για να πραγματοποιήσουν ερωτήματα προς το μητρώο.
- *rdqlQuery*. Η μέθοδος αυτή χρησιμοποιεί τη γλώσσα ερωτημάτων RDQL για τη πραγματοποίηση ερωτήσεων προς το μητρώο Grimoires.

4.4.2.6 Η υπηρεσία Διαμεσολάβησης

Η υπηρεσία διαμεσολάβησης (Resource Broker) χρησιμοποιείται μέσω της διεπαφής Selection για την επιλογή υποστάσεων κάθε αφηρημένη διεργασία μιας ροής εργασίας. Εκτελεί την επιλογή με τη χρησιμοποίηση διαμεσολάβησης βασισμένη στη διαπραγμάτευση συμβολαίων SLA (βλ. ενότητα 2.4.6). Ο στόχος της προσέγγισης είναι η επιλογή να ολοκληρώνεται με τρόπο που να ικανοποιεί τις απαιτήσεις και τους περιορισμούς και των φορέων παροχής υπηρεσιών και των πελατών.

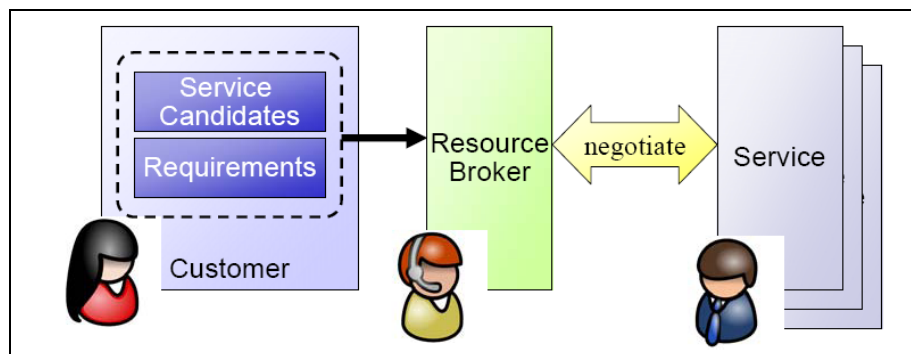
Η συγκεκριμένη υπηρεσία κρύβει από το χρήστη τον τρόπο διαχείρισης των πόρων και τις λεπτομέρειες που τους αφορούν. Επίσης εξειδικεύεται αναλόγως των πόρων που χειρίζεται και για να κάνει λογικές επιλογές πρέπει να γνωρίζει τι είδους πόρους διαχειρίζεται και σε τι κατάσταση είναι αυτοί. Ο διαμεσολαβητής μπορεί να εξειδικεύεται αναλόγως και με την οντότητα που παρέχει τις υπηρεσίες, καθώς αυτή η οντότητα μπορεί να ακολουθεί διαφορετικές πολιτικές σε θέματα όπως το κόστος.

Τα παραπάνω περιγράφονται στο σχήμα της επόμενης σελίδας (Σχήμα 28) όπου παρά την ύπαρξη μόνο ενός φορέα παροχής υπηρεσιών στο σχήμα, ο διαμεσολαβητής προσπαθεί συνήθως να διαπραγματευτεί συμβόλαια SLA με παραπάνω από έναν προμηθευτές. Επιπλέον στο σχήμα φαίνεται ότι ο διαμεσολαβητής είναι εξωτερική υπηρεσία και επομένως μπορεί να προσφέρεται και από τρίτους.



Σχήμα 28: Μοντέλο διαμεσολάβησης

Όπως φαίνεται στο σχήμα που ακολουθεί (Σχήμα 29), η διαδικασία διαμεσολάβησης αρχικοποιείται από τον πελάτη υπηρεσιών, ο πελάτης παρέχει έναν κατάλογο υποψηφίων φορέων παροχής υπηρεσιών καθώς επίσης και ένα σύνολο απαιτήσεων με όρους SLA (βλ. ενότητα 2.4.6.5) στο διαμεσολαβητή. Οι προαναφερθείσες απαιτήσεις εκφράζουν τις απαιτήσεις των πελατών σε όρους ποιότητας υπηρεσίας και επομένως οι υπηρεσίες που επιλέγονται από το διαμεσολαβητή θα πρέπει να ικανοποιούν τις απαιτήσεις των πελατών ενώ στην περίπτωση που ο φορέας παροχής υπηρεσιών επιτρέπει την επιλογή του εύρους τιμών για τους ιδιαίτερους όρους SLA, ο διαμεσολαβητής μπορεί να υπολογίσει τις τιμές που ικανοποιούν και τις απαιτήσεις του πελάτη και τους περιορισμούς του προμηθευτή.

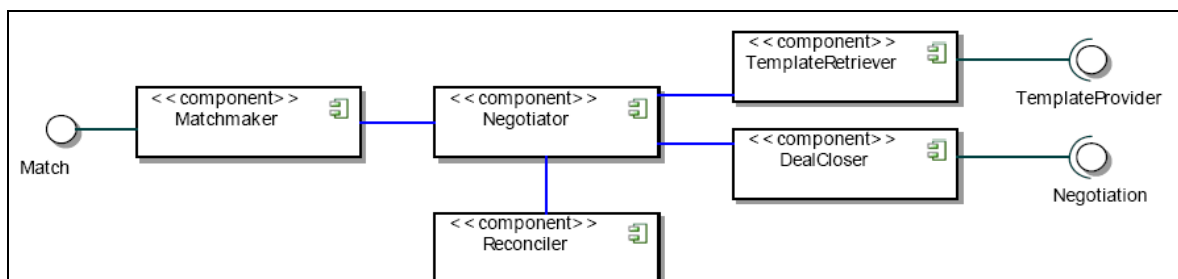


Σχήμα 29: Ανταλλαγή δεδομένων μεταξύ καταναλωτή και διαμεσολαβητή

Στο σχήμα της επόμενης σελίδας (Σχήμα 30) παρουσιάζεται η εσωτερική αρχιτεκτονική του διαμεσολαβητή που αποτελείται από πέντε (5) στοιχεία όπως αυτά αναλύονται στη συνέχεια:

Η σελίδα αυτή είναι σκόπιμα λευκή

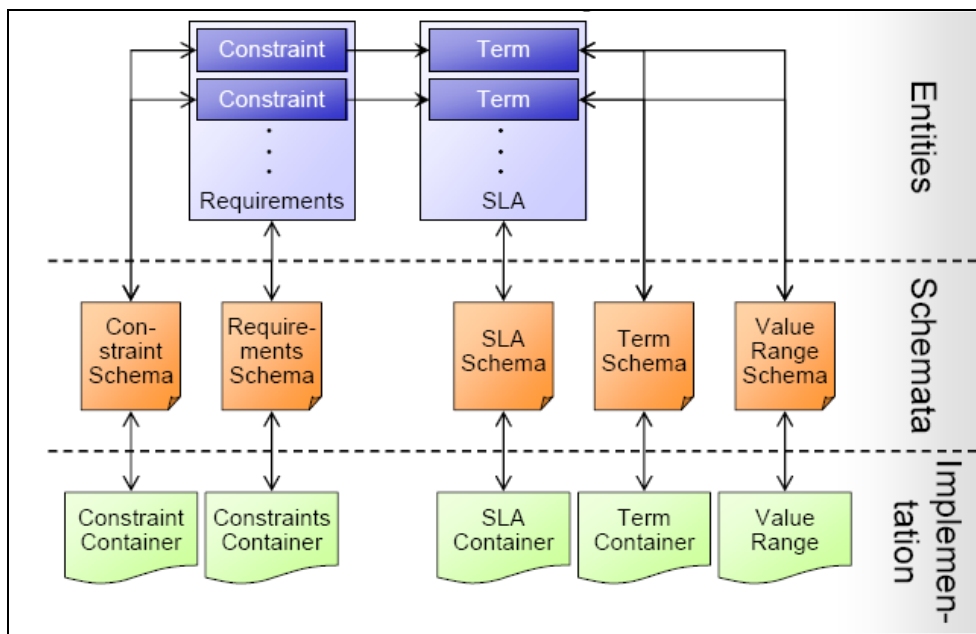
- Matchmaker: Εφαρμόζει το πρωτόκολλο που χρησιμοποιείται για την επικοινωνία με τον πελάτη και καθορίζει τις δομές δεδομένων που χρησιμοποιούνται για να εκφράσουν τις απαιτήσεις του πελάτη.
- Negotiator: Ενορχηστρώνει όλα τα άλλα συστατικά. Αρχικοποιείται από τον Matchmaker και προχωρεί τη διαδικασία διαπραγμάτευσης μέχρι την επίτευξη συμφωνίας SLA με έναν προμηθευτή ή ολοκλήρωσης της διαδικασίας.
- Reconciler: Συσχετίζει τις απαιτήσεις των πελατών με τους όρους των πρότυπων προσφερόμενων συμβολαίων SLA και αξιολογεί το βαθμό ικανοποίησης των όρων με τις απαιτήσεις. Με βάση αυτήν την αξιολόγηση μπορεί (υπό έλεγχο του Negotiator) να αλλάξει όρους που δεν ικανοποιούν τις απαιτήσεις με όρους που τις ικανοποιούν.
- TemplateRetriever και DealCloser: Είναι μηχανές πρωτοκόλλων που χρησιμοποιούνται για την επικοινωνία με τον φορέα παροχής υπηρεσιών. Ο TemplateRetriever παίρνει τα πρότυπα SLA που προσφέρονται από έναν δεδομένο φορέα παροχής υπηρεσιών μέσω της διαδικασίας WS-Agreement για την ανάγνωση των προτύπων από το WS-ResourceProperties. Ο DealCloser ζητά από τον προμηθευτή να δεχτεί την προσφορά που πρέπει να είναι η ίδια με ένα από τα προσφερόμενα πρότυπα με χρήση της WS-Agreement.



Σχήμα 30: Εσωτερική αρχιτεκτονική διαμεσολαβητή

Η σελίδα αυτή είναι σκόπιμα λευκή

Το μοντέλο δεδομένων (Data Model) που περιγράφεται ακολούθως είναι μια αφαίρεση των οντοτήτων που χρησιμοποιούνται από το διαμεσολαβητή. Στο σχήμα της επόμενης σελίδας (Σχήμα 31) παρουσιάζονται οι σχέσεις μεταξύ των διαμορφωμένων οντοτήτων, τα σχήματα που χρησιμοποιούνται για να περιγράψουν τις οντότητες και τους αφηρημένους υποδοχείς (containers) που αντιπροσωπεύουν τις οντότητες στο επίπεδο εφαρμογής.



Σχήμα 31: Μοντέλο δεδομένων

Τα σχήματα που παρουσιάζονται στο παραπάνω σχήμα (Σχήμα 31) αντιπροσωπεύουν τις εννοιολογικές τους κατηγορίες. Κάθε εννοιολογικό σχήμα μπορεί να πραγματοποιηθεί από διάφορα πραγματικά σχήματα. Παραδείγματος χάριν, το «SLA Schema» θα μπορούσε να πραγματοποιηθεί από το σχήμα WS-Agreement ενώ επίσης θα πρέπει να σημειωθεί ότι και τα πέντε σχήματα είναι εννοιολογικά ανεξάρτητα.

Οι διαμορφωμένες οντότητες είναι τα αντικείμενα που ο μεσολαβητής καθώς και ο πελάτης και ο φορέας παροχής υπηρεσιών χρησιμοποιούν κατά την επικοινωνία τους και διαμορφώνονται από το στοιχείο που είναι σχηματοποιημένο σύμφωνα με τα αντίστοιχα

Η σελίδα αυτή είναι σκόπιμα λευκή

σχήματα. Αυτά τα σχήματα επομένως περιγράφουν το στοιχείο που είναι εξωτερικά ορατό.

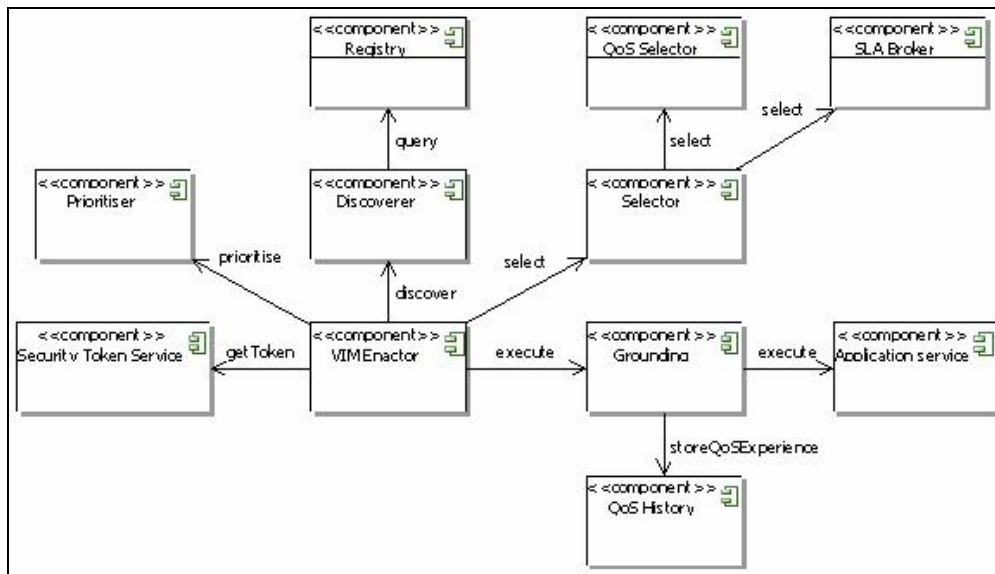
Βάσει όλων των παραπάνω ο διαμεσολαβητής εκτελεί τα παρακάτω βήματα για την επιλογή των διεργασιών:

- Λήψη των σημείων με τα οποία θα γίνει η διαπραγμάτευση συμβολαίων SLA: Λαμβάνει μια αφηρημένη υπηρεσία από την οποία εξάγει τις υποψήφιες διεργασίες για να υλοποιήσουν την αφηρημένη υπηρεσία, όπως αυτές έχουν βρεθεί μέσω της διαδικασίας της εύρεσης (Discovery). Από κάθε υποψήφια στη συνέχεια αποσπά μια αναφορά προς το τελικό σημείο με το οποίο θα γίνει η διαπραγμάτευση SLA.
- Λήψη των απαιτήσεων του χρήστη: Ο καταναλωτής της υπηρεσίας έχει συγκεκριμένες απαιτήσεις για τις οποίες θα γίνει διαπραγμάτευση σε όρους SLA. Ο επιλογέας εξάγει τις πληροφορίες από το αντικείμενο της αφηρημένης διεργασίας.
- Κλήση του διαμεσολαβητή. Ο διαμεσολαβητής διαπραγματεύεται τα συμβόλαια SLA με τις υποψήφιες υπηρεσίες λαμβάνοντας υπόψη τις απαιτήσεις που λαμβάνει από τους διαπραγματευόμενους. Εάν η διαπραγμάτευση είναι επιτυχής, ο επιλεγμένος υποψήφιος και η αντίστοιχη SLA επιστρέφονται.
- Μετάφραση του αποτελέσματος. Το αποτέλεσμα που επιστρέφεται από το διαμεσολαβητή μεταφράζεται στο περιβάλλον σημασιολογικής περιγραφής υπηρεσιών. Η υποψήφια υπηρεσία που επιλέγεται αντιπροσωπεύεται από την αντίστοιχη αναφορά προς το τελικό σημείο διαπραγμάτευσης της SLA. Η αναφορά αυτή συνδέεται με την αντίστοιχη συμπαγή υπηρεσία και η SLA που προέκυψε ως αποτέλεσμα της διαπραγμάτευσης συνδέεται με τη μορφή ειδικών σχολίων. Ως τελικό βήμα συνδέεται η επιλεγμένη υπηρεσία με την αφηρημένη διεργασία και την επιστρέφεται στον enactor.

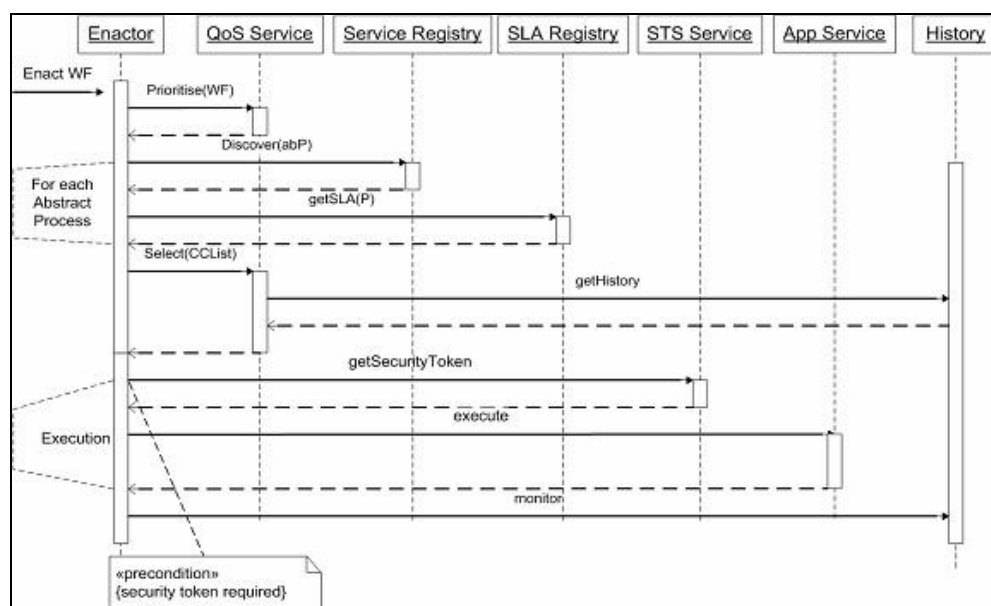
Η σελίδα αυτή είναι σκόπιμα λευκή

4.4.3 Αρχιτεκτονική Σχεδίαση του Μοντέλου Εικονικής Υποδομής

Βάσει όσων έχουν περιγραφεί στις προηγούμενες ενότητες για το Μοντέλο Εικονικής Υποδομής (Grid Virtual Infrastructure Model), στα παρακάτω σχήματα παρουσιάζεται η αρχιτεκτονική του μοντέλου και οι αλληλεπιδράσεις των υπηρεσιών τόσο σε μορφή αρχιτεκτονικού διαγράμματος (Σχήμα 32) όσο και σε μορφή αλλαγής πληροφοριών και καταστάσεων – Sequence Diagram (Σχήμα 33):



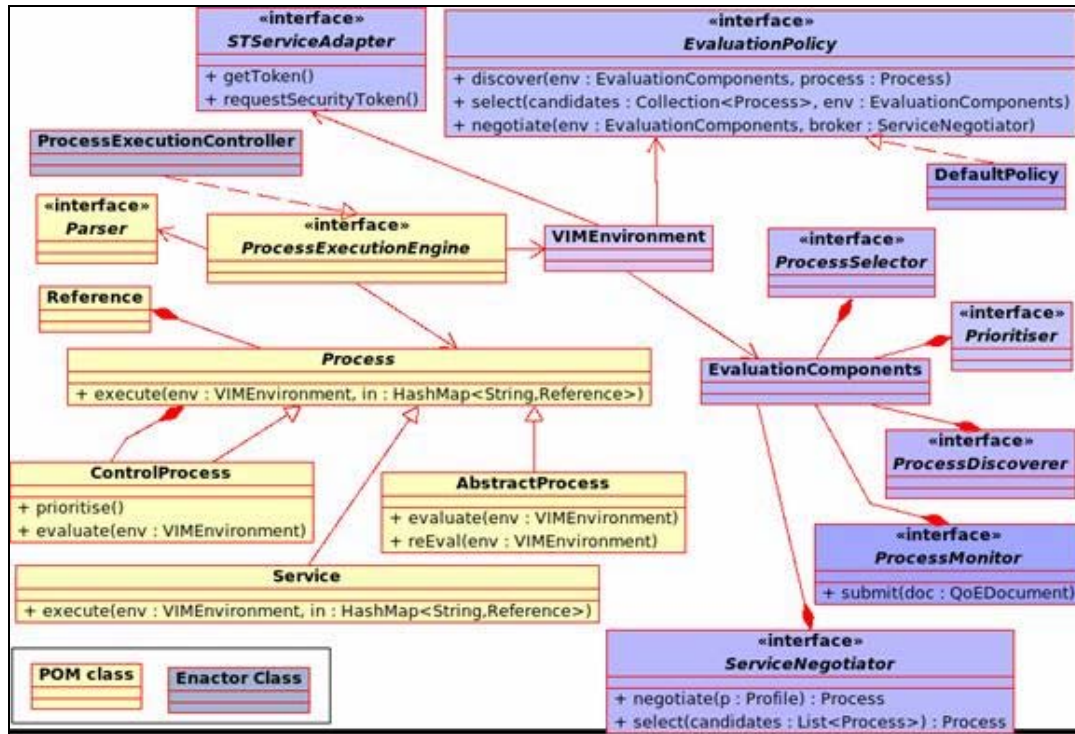
Σχήμα 32: Αρχιτεκτονική του μοντέλου εικονικής υποδομής και αλληλεπιδράσεις



Σχήμα 33: Διάγραμμα αλληλεπιδράσεων

Η σελίδα αυτή είναι σκόπιμα λευκή

Τέλος, στο σχήμα που ακολουθεί (Class Diagram – Σχήμα 34) παρουσιάζεται η υλοποίηση των παραπάνω στοιχείων και οι διεπαφές αυτών για την επίτευξη των αλληλεπιδράσεων μεταξύ τους.



Σχήμα 34: Σχηματική απεικόνιση υλοποίησης διεπαφών και αλληλεπιδράσεων

Η σελίδα αυτή είναι σκόπιμα λευκή

5

Ποιότητα Υπηρεσίας

Στο συγκεκριμένο κεφάλαιο παρουσιάζεται η έννοια της Ποιότητας Υπηρεσίας καθώς και η ανάγκη διασφάλισης αυτής για την υιοθέτηση επιχειρηματικών μοντέλων σε περιβάλλον πλέγματος. Επιπλέον, περιγράφονται οι παράμετροι που καθορίζουν και «ποσοτικοποιούν» την ποιότητα και οι κατηγορίες απαιτήσεων ποιότητας για συγκεκριμένες παραμέτρους. Το κεφάλαιο καταλήγει αναφέροντας μοντέλα δρομολόγησης εργασιών για την εκτίμηση και διασφάλιση της προσφερόμενης ποιότητας.

5.1 Γενικά Στοιχεία

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 παράγοντας αξιολόγησης.

Η μελέτη μηχανισμών απεικόνισης ροών εργασίας με στόχο το καθορισμό συγκεκριμένης ροής εργασίας απαιτεί την ανάλυση θεμάτων αναφορικά με την ποιότητα υπηρεσίας που θα

προσφέρεται για το σύνολο της ροής εργασίας καθώς η επιλογή των πόρων και των παρόχων θα πρέπει να έχει ως αποτέλεσμα παροχή υπηρεσιών με συγκεκριμένο χρόνο απόκρισης, απόδοση, διαθεσιμότητα και ασφάλεια, έτσι ώστε η παραγωγικότητα και η χρησιμότητα των συνδυασμένων πόρων να είναι πολύ μεγαλύτερη από τις δυνατότητές τους ξεχωριστά. Στο συγκεκριμένο κεφάλαιο διατυπώνονται θέματα σχετικά με τις απαιτήσεις και τις παραμέτρους που πρέπει να ικανοποιούνται προκειμένου να υπάρχει δυνατότητα δυναμικής απεικόνισης ροών εργασίας για διαφορετικές επιχειρηματικές διαδικασίες. Επίσης, αυτές οι παράμετροι ταξινομούνται με βάση την προέλευσή τους (για παράδειγμα, μπορεί να πρόκειται για παραμέτρους της εφαρμογής ή για παραμέτρους που ορίζονται από το χρήστη) ενώ αναλύονται συγκεκριμένες περιπτώσεις για τον προαναφερθέν μηχανισμό απεικόνισης ροών εργασίας.

5.1.1 Η Ποιότητα Υπηρεσίας σε περιβάλλον Πλέγματος

Η σημασία της ποιότητας των παρερχομένων υπηρεσιών σε περιβάλλον πλέγματος σύστημα είναι πολύ μεγάλη, αφού αυτό που κάνει μια υπηρεσία προσιτή 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 (Service Oriented Architecture – βλ. ενότητα 4.2) έκανε πιο εμφανή την ανάγκη προσδιορισμού της ποιότητας των υπηρεσιών. Η συμμετοχή πολλών διαφορετικών συστημάτων στο ίδιο περιβάλλον πλέγματος απαιτεί μια κοινή παράμετρο αξιολόγησής τους αλλά και των υπηρεσιών που προσφέρουν. Επιπρόσθετα, η αρχιτεκτονική που χρησιμοποιείται σε τέτοια περιβάλλοντα είναι βασισμένη στις υπηρεσίες και λόγω αυτού, η εισαγωγή του παράγοντα της προσφερόμενης ποιότητας υπηρεσίας είναι εφικτή σε όλα τα επίπεδα της δομής του πλέγματος.

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 του χρήστη.

Υπάρχουν αρκετές παράμετροι που εξαιτίας της ιδιομορφίας του περιβάλλοντος πλέγματος (δυναμική εισαγωγή ετερογενών υπολογιστικών και μη συστημάτων) δεν είναι εύκολο να προσδιοριστούν, καθώς βασικές πληροφορίες μπορούν να εξαχθούν από προγράμματα παρακολούθησης (monitoring) και άλλα συστήματα μοντέλα προτυποποίησης των πόρων. Βάσει των απαιτήσεων ενός χρήστη σε περιβάλλον πλέγματος και των πληροφοριών που μπορούν να εξαχθούν από το περιβάλλον πλέγματος, παρουσιάζονται οι παρακάτω κατηγορίες πληροφορίας αναφορικά με την ποιότητα υπηρεσίας:

- Διαθεσιμότητα: η πληροφορία αυτή μπορεί να δοθεί με μορφή ένδειξης, για το αν είναι διαθέσιμο το στιγμιότυπο της υπηρεσίας (service up-and-running), και άλλα απαραίτητα αρχεία και βιβλιοθήκες.
- Απόδοση: ο προσδιορισμός αυτής της παραμέτρου είναι αρκετά δύσκολος εξαιτίας της αδυναμίας μοντελοποίησης ενός γενικού τρόπου εξαγωγής της πληροφορίας. Παρόλα αυτά μπορούμε να θεωρήσουμε ότι υπολογίζοντας το χρόνο απόκρισης, τη ταχύτητα του επεξεργαστή και κάποιες άλλες τιμές ενδεικτικές για την αποδοτικότητα και ικανότητα του συστήματος, έχουμε μια αφηρημένη εικόνα για την παράμετρο αυτή.
- Αξιοπιστία: η παράμετρος αυτή πηγάζει από την επεξεργασία των αποτελεσμάτων της διαθεσιμότητας. Η διαθεσιμότητα μιας υπηρεσίας για μεγάλο διάστημα χρόνου, της αποδίδει αξιοπιστία.

- Κόστος: η παράμετρος αυτή έχει να κάνει αποκλειστικά με την τιμολόγηση των υπηρεσιών σε σχέση με το χρόνο κλπ.
- Στατιστικά στοιχεία για την ποιότητα του αποτελέσματος (Quality of the Result): τα στοιχεία αυτά επιτυγχάνονται από την αξιολόγηση των υπηρεσιών από τους πελάτες, αποδίδοντας ένα βαθμό ικανοποίησης του χρήστη.
- Στατιστικά στοιχεία για την παραβίαση των συμβολαίων SLA: μια ένδειξη για την κατάσταση μιας υπηρεσίας σε σχέση με το συμφωνηθέν SLA θα αποτελούσε χρήσιμη πληροφορία

Οι δύο πρώτες κατηγορίες ανήκουν στα ποσοτικά χαρακτηριστικά, σύμφωνα με τον αρχικό διαχωρισμό, ενώ τα υπόλοιπα αφορούν την ποιότητα της παρεχόμενης υπηρεσίας και ανήκουν στα ποιοτικά. Παρόλο που οι ιδιότητες οι σχετικές με την ποιότητα της υπηρεσίας (ποιοτικά χαρακτηριστικά) είναι πολύ σημαντικές στην αξιολόγηση της προσφερόμενης ποιότητας, είναι δύσκολο να υπολογιστούν αντικειμενικά, αφού βασίζονται κατά πολύ στην ανάδραση πληροφορίας από τον χρήστη. Λόγω αυτού οι μηχανισμοί απεικόνισης ροών εργασίας επικεντρώνονται σε ποσοτικά χαρακτηριστικά όπως η διαθεσιμότητα (availability), το κόστος (cost) και ο χρόνος εκτέλεσης / μιας εργασίας (execution time).

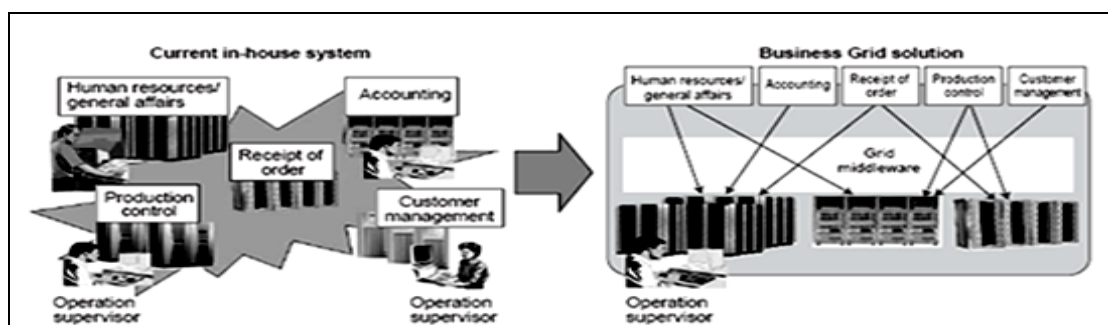
5.1.2 Από το Ακαδημαϊκό στο Επιχειρηματικό Πλέγμα

Η χρήση συμβολαίων παροχής υπηρεσιών (SLAs) στις υπηρεσίες έδωσε τις προδιαγραφές για την ανάπτυξη της τεχνολογίας πλέγματος και εκτός του ερευνητικού τομέα. Στο χώρο του e-business οι απροσδόκητες αλλαγές στο φόρτο εργασιών, στον όγκο πληροφοριών ή στις απαιτήσεις επεξεργασίας δημιούργησαν την απαίτηση για ευέλικτα, μικρού κόστους και μεγάλης αξιοπιστίας συστήματα. Πάνω σε αυτήν την απαίτηση στηρίχθηκε η τεχνολογία πλέγματος για την δημιουργία του επιχειρηματικού (Business) Grid [47] με σκοπό την

εμπορική εκμετάλλευση των παρεχόμενων υπηρεσιών. Πιο συγκεκριμένα οι στόχοι του επιχειρηματικού πλέγματος είναι οι ακόλουθοι:

- Κοινόχρηστη χρήση πόρων, μέσω δικτύωσης υπολογιστών των υπολογιστών απομακρυσμένων κέντρων πληροφοριών (data centers)
- Μείωση του κόστους στις επιχειρήσεις με τη χρήση ευέλικτων υποδομών και αυτόνομων συστημάτων ελέγχου
- Δυναμική εισαγωγή και χρήση πόρων με την ταυτόχρονη διοχέτευση φόρτου εργασίας σε διαφορετικές απομακρυσμένες περιοχές
- Απλοποίηση, συνολικά των υποδομών και τη δημιουργία αξιόπιστων συστημάτων που θα υποστηρίζουν διασπορά εργασίας, ανάκτηση κατεστραμμένων πληροφοριών και άλλα λειτουργικά συστήματα και συστήματα ασφαλείας.

Τα παραπάνω παρουσιάζονται γραφικά στο σχήμα της επόμενης σελίδας (Σχήμα 35), όπου όπως φαίνεται η υποδομή του επιχειρηματικού πλέγματος βελτιστοποιεί την εκμετάλλευση των πόρων μιας επιχείρησης με αποτέλεσμα τη μείωση του κόστους και την αύξηση της παραγωγικότητας.



Σχήμα 35: Από το ακαδημαϊκό στο επιχειρηματικό πλέγμα

5.1.3 Προσέγγιση Λειτουργίας του Επιχειρηματικού Μοντέλου του Πλέγματος

Η χρήση ενός περιβάλλοντος πλέγματος δεν περιορίζεται στην εκμετάλλευση των παρεχόμενων συστημάτων εντός μιας επιχείρησης με σκοπό την μείωση κόστους, την

Η σελίδα αυτή είναι σκόπιμα λευκή

αύξηση της παραγωγικότητας και μείωση του κύκλου παραγωγής, αλλά και στην εμπορική εκμετάλλευση αυτών των τεχνολογιών και των υπηρεσιών πλέγματος. Η απαίτηση της αγοράς για εξειδίκευση στην παροχή υπηρεσιών και η συγκεκριμένη υπηρεσιοστρεφής αρχιτεκτονική του πλέγματος (βλ. ενότητα 4.2) επιτρέπει το προαναφερθέν.

Για την περιγραφή της λειτουργίας ενός τέτοιου συστήματος κρίνεται απαραίτητη η αναφορά των ρόλων και των οντότητων που συμμετέχουν σε αυτό. Βάσει αυτού, αναφέρεται αρχικά ο πάροχος της υπηρεσίας (Service Provider). Πρόκειται για γενικό όρο και μπορεί να διασπαστεί και σε άλλες υπό-οντότητες, αφού οι παρεχόμενες υπηρεσίες που σχετίζονται είναι υπηρεσίες δικτύου, πόρων (αποθηκευτικών, υπολογιστικών κλπ.), εφαρμογών (application services) κ.α. Κάποιες από τις παραπάνω υπηρεσίες μπορούν να παρέχονται από τον ίδιο οργανισμό ή να είναι διαμοιρασμένες σε περισσότερους. Σε κάθε περίπτωση, για την ομαλή και αποδοτική συνεργασία μεταξύ των πάροχων και των χρηστών μιας υπηρεσίας υπάρχουν τα συμβόλαια παροχής υπηρεσιών (SLAs) που διασφαλίζουν τους όρους συνεργασίας. Όπως αναφέρθηκε και σε προηγούμενα κεφάλαια, τα SLAs θα πρέπει να διασφαλίσουν και την ποιότητα των παρεχομένων υπηρεσιών και αυτό μπορεί να επιτευχθεί με την εισαγωγή στο σύστημα και ενός άλλου ρόλου, του QoS Provider, που αναφέρεται σε μια οντότητα η οποία θα επικοινωνεί με άλλους πάροχους υπηρεσιών, θα συλλέγει πληροφορίες και θα προσδιορίζει την ποιότητα των προσφερομένων υπηρεσιών από και μεταξύ κάθε πλευράς. Ο ρόλος αυτός είναι πολύ σημαντικός αφού διασφαλίζει την ποιότητα της υπηρεσίας πλέγματος που παρέχεται. Φυσικά η τελευταία οντότητα του συστήματος είναι ο πελάτης / χρήστης των παρεχομένων υπηρεσιών του πλέγματος, από τις απαιτήσεις του οποίου γίνονται όλες οι προσπάθειες για τη βελτίωση της ποιότητας των υπηρεσιών.

5.1.3.1 QoS Provider

Σε συνέχεια των παραπάνω, στην παρούσα παράγραφο αναλύονται και περιγράφονται οι προδιαγραφές λειτουργίας του QoS Provider. Η συγκεκριμένη οντότητα (ως πάροχος QoS πληροφορίας) θα πρέπει:

- Να έχει πρόσβαση σε πληροφορίες των πόρων μιας υποδομής πλέγματος που άλλοι χρήστες μπορεί να μην έχουν
- Να έχει υλοποίηση γενικής χρήσης έτσι ώστε να μπορεί να εφαρμοστεί σε διαφόρων τύπων πόρους
- Να είναι συμβατή με τις τεχνολογίες που εφαρμόζονται στα περιβάλλοντα πλέγματος
- Η υλοποίησή της να είναι δημοσίως διαθέσιμη ώστε οποιοσδήποτε επιθυμεί on-demand πληροφορία να μπορεί να την αποκτήσει
- Να γίνεται καταγραφή και επεξεργασία των πληροφοριών αυτών όχι μόνο για τεκμηρίωση της ποιότητας των προσφερομένων υπηρεσιών αλλά και για βελτίωση της λειτουργίας όλου του πλέγματος.

Η ποιότητα υπηρεσιών σε περιβάλλον πλέγματος σχετίζεται άμεσα με τους όρους που την προσδιορίζουν και την καθορίζουν. Η παραμετροποίηση αυτών των όρων και η εξαγωγή της αντίστοιχης πληροφορίας βασίζεται στις παραμέτρους που αναφέρθηκαν σε προηγούμενη ενότητα (ενότητα 5.3). Στη συνέχεια αναλύουμε τους όρους που χρησιμοποιούνται από το μηχανισμό απεικόνισης ροών εργασίας για την επιλογή υπηρεσιών πλέγματος που προσφέρουν το επιθυμητό επίπεδο ποιότητας υπηρεσίας με βάση τις απαιτήσεις του χρήστη όπως αυτές εκφράζονται με συγκεκριμένους όρους. Είναι προφανές ότι υπάρχουν πολυάριθμες παράμετροι που είναι ιδιαίτερου ενδιαφέροντος στις υποδομές πλεγμάτων. Η μελέτη εστιάστηκε αρχικά στη Διαθεσιμότητα, στην Αποδοτικότητα, στην Αξιοπιστία και στο Κόστος, καθώς έχουν επιλεγεί ως αντιπροσωπευτικές ενδεικτικές παράμετροι για τη

συνολική διαδικασία απεικόνισης ροών εργασίας και το σύνολο των παραμέτρων μπορεί να επεκταθεί ή να τροποποιηθεί ώστε να συμπεριλάβει άλλες παραμέτρους.

5.1.3.2 Διαθεσιμότητα

Ο όρος διαθεσιμότητα (Availability) περιλαμβάνει παραμέτρους όπως τη διαθεσιμότητα των εφαρμογών, των πόρων, του δικτύου κ.α. Το σημαντικότερο στοιχείο στην ποιότητα μιας υπηρεσίας πλέγματος, αλλά και γενικά μιας οποιαδήποτε υπηρεσίας, είναι το κατά πόσο αυτή και οι τυχόν εξαρτήσεις αυτής είναι διαθέσιμες. Ποιο αναλυτικά τα στοιχεία που χαρακτηρίζουν τη συγκεκριμένη παράμετρο είναι τα ακόλουθα:

- Διαθέσιμος ελεύθερος χώρος
- Εύρος δικτύου
- Απαραίτητες βιβλιοθήκες-αρχεία
- Υποστηριζόμενα πρωτόκολλα υπηρεσιών
- Κατάσταση υπηρεσιών εφαρμογών

Τα παραπάνω είναι οι στοιχειώδεις προϋποθέσεις για την εκτέλεση μιας εφαρμογής σε περιβάλλον πλέγματος. Επιπρόσθετα, θέτοντας σαν στόχο τη χρήση πληροφορίας για τον καθορισμό της διαθεσιμότητας, δε γίνεται αναζήτηση συγκεκριμένων παραμέτρων για κάθε είδος συστήματος ή πόρου.

5.1.3.3 Αποδοτικότητα

Ένα επιπλέον χαρακτηριστικό που είναι πολύ σημαντικό και καθοριστικό για την ποιότητα της υπηρεσίας που παρέχεται είναι η απόδοση των συστημάτων (Performance) που συμμετέχουν στο πλέγμα. Ο καθορισμός τέτοιων παραμέτρων που θα προσδιορίζουν την απόδοση, είναι αρκετά δύσκολος και διαφέρει από σύστημα σε σύστημα. Επίσης η ταχύτητα με την οποία μπορεί να αλλάξει η κατάσταση ενός συστήματος από αποδεκτής απόδοσης σε μη αποδεκτής δυσκολεύει ακόμη περισσότερο την μοντελοποίηση τέτοιων μηχανισμών.

Λαμβάνοντας υπόψη ότι το μεγαλύτερο μέρος των συστημάτων που συμμετέχουν θα είναι υπολογιστικές μονάδες, κάποιες μελέτες έχουν καταλήξει σε παραμέτρους όπως :

- Συχνότητα ρολογιού επεξεργαστή
- Υπολογισμός MFLOPs (Millions Floating Operations per second)
- Μέσο Εύρος ζώνης μνήμης
- Μέσο εύρος ζώνης σκληρού δίσκου
- Διαθέσιμη μνήμη
- Χρόνος εκτέλεσης εργασίας

Καθώς μια σειρά παραμέτρων μπορούν να προστεθούν στις παραπάνω, η μελέτη θα πρέπει να επικεντρωθεί σε αυτήν (ή αυτές) που αποτελούν βασικά κριτήρια. Βάσει αυτού ως χαρακτηριστικό της αποδοτικότητας ενός κόμβου θεωρούμε το χρόνο που απαιτείται για την εκτέλεση μιας εργασίας καθώς αυτός είναι αντιπροσωπευτικός και περιλαμβάνει άλλα χαρακτηριστικά όπως η διαθέσιμη μνήμη και η συχνότητα ρολογιού του επεξεργαστή.

5.1.3.4 Αξιοπιστία

Η αξιοπιστία (Reliability) κάθε παρεχόμενης υπηρεσίας έχει άμεση σχέση με την διαθεσιμότητά της σε σύγκριση με την χρονική διάρκεια αυτής. Εκτός από αυτό όμως, η αξιοπιστία εξαρτάται και από παραμέτρους όπως η συχνότητα εμφάνισης σφαλμάτων ή την πιθανότητα εμφάνισής τους.

Συνολικά όμως η αξιοπιστία της υπηρεσίας διαπιστώνεται από μακροχρόνια παρακολούθηση των παραμέτρων της διαθεσιμότητας, απόδοσης και συνολικής επεξεργασίας τους.

5.1.3.5 Κόστος

Το κόστος (Cost) αναφέρεται στο κόστος που απαιτείται για την ολοκλήρωση της εκτέλεσης μιας υπηρεσίας. Θα πρέπει στο σημείο αυτό να σημειωθεί ότι η εκτέλεση δεν αναφέρεται

υποχρεωτικά σε εκτέλεση εργασιών αλλά σε υπηρεσίες (για παράδειγμα αίτηση και δέσμευση αποθηκευτικού χώρου σε κάποιο πόρο του πλέγματος).

Επιπρόσθετα εκτός από το κόστος μιας υπηρεσίας στο περιβάλλον πλέγματος υφίσταται και το κόστος ως ρήτρα που μπορεί κάποιο από τα εμπλεκόμενα μέρη να πρέπει να πληρώσει στην περίπτωση μη εκπλήρωσης των συμβατικών του υποχρεώσεων – όπως αυτές περιγράφονται στα συμβόλαια παροχής υπηρεσιών (SLAs).

5.2 Παρούσα κατάσταση - Σχετικές εργασίες

Όπως έχει ήδη αναφερθεί, ο όρος πλέγμα αναφέρεται σε μεγάλης κλίμακας διαχείριση και κοινή χρήση πόρων. Η συνεργασία των παρόχων υπηρεσιών και των χρηστών είναι απαραίτητο να καθορίζεται από συγκεκριμένους κανόνες, έτσι ώστε να είναι ξεκάθαρη η σχέση μεταξύ παροχών υπηρεσιών και καταναλωτών, ποιος επιτρέπεται να χρησιμοποιήσει τους πόρους, ποιους πόρους, κάτω από ποιες συνθήκες κλπ. Οι κανονισμοί αυτοί προτυποποιήθηκαν και έτσι δημιουργήθηκαν τα συμβόλαια SLA (βλ. ενότητα 2.4.6).

Οι παλαιότερες προσπάθειες για την εισαγωγή της έννοιας της ποιότητας υπηρεσίας στη διαχείριση των πόρων είχαν μεγάλο ενδιαφέρον και σημαντικά αποτελέσματα. Υπάρχουν διάφορες παράμετροι για τους μηχανισμούς διαχείρισης των παραμέτρων της ποιότητας υπηρεσίας προκειμένου να βελτιστοποιηθεί η διαδικασία επιλογής στα πλαίσια μιας ροής εργασίας. Γενικά, ο τρόπος που γίνεται αντιληπτή η λειτουργία QoS στις τρέχουσες αρχιτεκτονικές πλεγμάτων είναι μέρος της διαδικασίας διαπραγμάτευσης SLA όπως περιγράφεται στις αναφορές [18], [19] και [20]. Οι περιορισμοί και οι προτιμήσεις του τελικού χρήστη αναλύονται συντακτικά σε αρκετούς παρόχους υπηρεσιών μέσω της λειτουργικότητας που παρέχεται από μεσίτη (συνήθως την υπηρεσία διαχείρισης SLA) για την ανεύρεση των κατάλληλων παρόχων υπηρεσιών. Ο Sahai [45] πρότεινε την υποστήριξη της QoS πληροφορίας σε SLA οντότητες στα ευρείας χρήσης (εμπορικά) πλέγματα. Αν και

πολύ ενδιαφέρουσα προσέγγιση, δεν αναπτύχθηκε σημαντικά και είναι ακόμα σε ερευνητικό στάδιο. Ένα γενικό μοντέλο με την ονομασία SNAP (Service Negotiation and Acquisition Protocol) προτάθηκε από τον Czajkowski [46]. Το μοντέλο αυτό διακρίνει τρία είδη συμβολαίων SLA (βλ. ενότητα 2.4.6) για την ανάθεση των εργασιών στο πλέγμα, Task SLA (TSLA), Resource SLA (RSLA), Bind SLA (BSLA). Το πρώτο από αυτά εμπεριέχει τις απαιτήσεις για την ποιότητα της υπηρεσίας που ζητήθηκε. Ο Keahey πρότεινε την αρχιτεκτονική VAS (Virtual Application Service) για την διαχείριση της ποιότητας στα πλέγματα. Η αρχιτεκτονική αυτή εισήγαγε καινούριους τρόπους διαπραγμάτευσης του επιπέδου της ποιότητα των υπηρεσιών και των απαιτήσεων. Βάσει των παραπάνω, η διαχείριση των πόρων του πλέγματος είναι σημαντική προκειμένου να επιτευχθεί βελτιστοποίηση της κατανομής τους [80], και κρίνεται καθοριστική ως προς τη δυνατότητα του πλέγματος να ανταποκριθεί στο απαιτούμενο επίπεδο ποιότητας υπηρεσιών προς όλους τους χρήστες [81], [82].

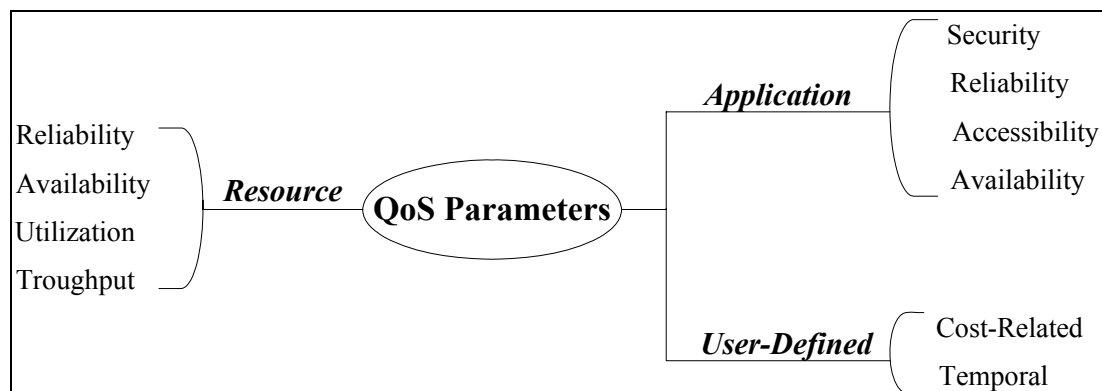
Τέλος, η προσέγγιση GARA (Globus Architecture for Reservation and Allocation) [21] αντιμετωπίζει το ζήτημα παροχής ποιότητας στο επίπεδο της διευκόλυνσης και της παροχής βασικών μηχανισμών για την υποστήριξη ποιότητας υπηρεσίας, δηλαδή διαμόρφωση, ανεύρεση, επιλογή και εκχώρηση πόρων. Η τρέχουσα έμφαση της παραπάνω προσέγγισης βρίσκεται στην υποστήριξη της ταυτοποίησης και έγκρισης αιτήματος εκχώρησης πόρων [76] με συγκεκριμένες παραμέτρους υπηρεσίας.

Στις επόμενες ενότητες παρουσιάζουμε μια καινοτόμο κατηγοριοποίηση των παραμέτρων που καθορίζουν και «ποσοτικοποιούν» την ποιότητα υπηρεσίας σε περιβάλλοντα πλέγματος καθώς οι παράμετροι αναμένεται να διαδραματίσουν σημαντικό ρόλο την εκτίμηση και την διασφάλιση της παρεχόμενης ποιότητας. Επιπρόσθετα, παρουσιάζονται οι κατηγορίες των απαιτήσεων ποιότητας βάσει των προαναφερθέντων παραμέτρων που αποτελούν βασικό

σημείο και προϋπόθεση επιτυχούς λειτουργίας του μηχανισμού που περιγράφεται στο έκτο κεφάλαιο της διατριβής.

5.3 Καινοτόμος Κατηγοριοποίηση Παραμέτρων Ποιότητας Υπηρεσίας

Βασικό σημείο για την επιλογή των υπηρεσιών / πόρων που θα εκτελέσουν μια εργασία σε περιβάλλον πλέγματος (όπως αυτή ορίζεται σε μια αφηρημένη ροή εργασίας) είναι οι παράμετροι που καθορίζουν την προσφερθείσα ποιότητα υπηρεσίας. Λόγω αυτού υπεισέρχεται το θέμα του ποιες είναι αυτές οι παράμετροι-περιορισμοί της ποιότητας υπηρεσίας, πώς καθορίζονται και πώς αυτές αντανakλούν τις προτιμήσεις του χρήστη, που θέλει να εκτελέσει μια εφαρμογή σε περιβάλλον πλέγματος. Υπάρχει πλήθος παραμέτρων ειδικού ενδιαφέροντος για τις αρχιτεκτονικές πλέγματος όπως αναφέρθηκε στην προηγούμενη παράγραφο. Οι παράμετροι μπορεί να καθορίζονται από την ίδια την εφαρμογή ή από το χρήστη ή μπορεί να είναι παράμετροι, που χαρακτηρίζουν τους πόρους / κόμβους του πλέγματος. Στο παρακάτω σχήμα παρουσιάζονται οι κατηγορίες των παραμέτρων οι οποίες αναλύονται στη συνέχεια (Σχήμα 36):



Σχήμα 36: Ταξινόμηση παραμέτρων ποιότητας υπηρεσίας

Η σελίδα αυτή είναι σκόπιμα λευκή

Οι παραπάνω κατηγορίες προκύπτουν ως συνέπεια της λογικής κατηγοριοποίησής τους, που σημαίνει ότι οι παράμετροι που ανήκουν στην ίδια κατηγορία έχουν κοινές ιδιότητες και είναι οι ακόλουθες:

- Παράμετροι εφαρμογής (application parameters). Σχετίζονται με την προσφερόμενη ποιότητα υπηρεσίας από την άποψη της εφαρμογής. Για παράδειγμα, ρυθμίζοντας ορισμένες παραμέτρους της εφαρμογής είναι δυνατόν να επηρεαστεί η διαθεσιμότητα (availability) του πόρου.
- Παράμετροι καθοριζόμενοι από χρήστη (user-defined parameters). Αποτελούν παραμέτρους που εκφράζουν τις απαιτήσεις / περιορισμούς του χρήστη που καταχωρεί το προς εκτέλεση application workflow. Συνήθως αποτελούν απαιτήσεις ή περιορισμούς σχετικά με το κόστος της όλης διαδικασίας ή σε σχέση με τη διαθεσιμότητα που πρέπει να προσφέρει ένας πόρος.
- Παράμετροι των πόρων (resource parameters). Αποτελούν τις παραμέτρους που σχετίζονται με όλα τα είδη πόρων, όπως υπολογιστικοί πόροι (computational resources), αποθηκευτικοί πόροι (storage resources) και πόροι δικτύου (network resources). Από την άποψη αυτή η διαδικτυακή υποδομή μπορεί να θεωρηθεί σαν ένα σύνολο αλληλεπιδρώντων πόρων που προσφέρουν ένα συγκεκριμένο επίπεδο ποιότητας υπηρεσίας.

5.4 Κατηγορίες απαιτήσεων

Εκτός των παραμέτρων και της ταξινόμησής τους, οι χρήστες μπορούν επίσης να δηλώσουν τους περιορισμούς και τις προτιμήσεις τους. Αυτή η πρωτοποριακή άποψη παρέχει επιπρόσθετη αξία στη διαδικασία απεικόνισης ροών εργασίας καθώς η επιλογή γίνεται λαμβάνοντας υπόψη και τις προτιμήσεις του χρήστη. Προκειμένου να είναι αυτό εφικτό,

προσδίδεται συντελεστής βαρύτητας στις παραμέτρους ποιότητας. Με βάση αυτό και θεωρώντας ως αρχικές ενδεικτικές παραμέτρους τη διαθεσιμότητα, το κόστος και το χρόνο εκτέλεσης προσδιορίζονται οι παρακάτω κατηγορίες απαιτήσεων.

5.4.1 Απαιτήσεις για το Κόστος

Όταν δημιουργείται η αφηρημένη ροή εργασίας, η υπηρεσία εύρεσης υπηρεσιών (βλ. ενότητα 0) αναλαμβάνει να βρει τις κατάλληλες υποστάσεις υπηρεσιών (service instances), που θα είναι υποψήφιες για να χρησιμοποιηθούν σε κάθε βήμα της αφηρημένης ροής εργασίας. Η προσφορά κάθε στιγμιότυπου υπηρεσίας χαρακτηρίζεται και από ένα συγκεκριμένο κόστος το οποίο λαμβάνεται υπόψη για την επιλογή του ή όχι. Το κόστος αυτό συνήθως εκφράζεται σε μία χρηματική ή νομισματική μονάδα (για παράδειγμα 10 Euro). Υπάρχουν διάφορα είδη απαιτήσεων σε σχέση με το κόστος. Θα πρέπει να αναφερθεί ότι αυτές οι απαιτήσεις θέτουν περιορισμό στο ποσό που θα ξοδευτεί, αλλά δεν αναφέρονται στο πώς θα γίνει διαχείριση του ποσού αυτού. Ο χρήστης μπορεί ίσως να ενδιαφέρεται για το πώς θα ξοδευτούν για παράδειγμα δε θα ήθελε να δώσει χρήματα για έναν πόρο τη στιγμή που υπάρχει αντίστοιχος πόρος που απαιτεί λιγότερα χρήματα. Σε αυτή την περίπτωση χρειάζεται μια επιπλέον παράμετρος η οποία θα συζητηθεί αργότερα. Ορισμένες κατηγορίες απαιτήσεων για το κόστος είναι:

- Απαίτηση ανώτερου συνολικού κόστους. Σε αυτή την περίπτωση ο χρήστης θέτει ένα ανώτατο όριο για το συνολικό κόστος το οποίο δεν πρέπει να ξεπεραστεί. Αν αθροιστούν τα κόστη, που χρειάζονται για τη χρήση των πόρων, δε θα πρέπει το ποσό που προκύπτει να είναι μεγαλύτερο από το ανώτατο αυτό όριο.
- Απαίτηση ανώτερου κόστους ανά υπηρεσία. Σε αυτή την περίπτωση ο χρήστης καθορίζει ένα ανώτατο όριο κόστους για κάθε τύπο υπηρεσίας. Το κόστος που

απαιτεί η χρήση των υπηρεσιών, που θα επιλεγθούν για την εκτέλεση των επιμέρους τύπων υπηρεσιών, δε θα πρέπει να ξεπερνάει το κόστος που έχει οριστεί για κάθε τύπο υπηρεσίας.

- Αδιαφορία για το ύψος του κόστους. Σε αυτή την περίπτωση ο χρήστης δεν ενδιαφέρεται για το πόσα χρήματα θα ξοδέψει. Η τρίτη αυτή περίπτωση ισοδυναμεί ουσιαστικά με τις προηγούμενες περιπτώσεις αλλά θέτοντας ως ανώτατο όριο κόστους ένα πάρα πολύ μεγάλο ποσό.
- Επιλογή της φθηνότερης λύσης. Σε αυτή την περίπτωση θα πρέπει να επιλέγονται για κάθε τύπο υπηρεσίας οι υπηρεσίες με το μικρότερο κόστος.

5.4.2 Απαιτήσεις για τη Διαθεσιμότητα

Όσον αφορά τη διαθεσιμότητα, αυτή υπολογίζεται σύμφωνα με πολύπλοκο μηχανισμό, που λαμβάνει υπόψη του στατιστικά στοιχεία σχετικά με τη χρήση ενός πόρου όπως αναφέρεται στην ενότητα 5.1.3.2. Η διαθεσιμότητα δίνεται με τη μορφή ποσοστού (για παράδειγμα 50% διαθεσιμότητα). Κατηγορίες απαιτήσεων που έχουν σχέση με τη διαθεσιμότητα (availability) είναι:

- Απαίτηση κατώτερου ορίου στη διαθεσιμότητα για κάθε υπηρεσία. Στην περίπτωση αυτή απαιτείται από τις προσφορές, που υλοποιούν κάθε μία τύπο υπηρεσίας, να έχουν διαθεσιμότητα η οποία δε θα είναι μικρότερη από το όριο που έχει τεθεί.
- Απαίτηση κατώτερου ορίου στη διαθεσιμότητα του συνδυασμού όλων των υπηρεσιών. Σε αυτή την περίπτωση απαιτείται η συνδυασμένη διαθεσιμότητα των υπηρεσιών, που υλοποιούν κάθε τύπο υπηρεσίας, να μην είναι κατώτερη από το όριο που έχει τεθεί.

- Αντίστοιχα και με τις κατηγορίες για το κόστος μπορεί να θεωρηθεί και εδώ η επιπλέον περίπτωση να μη τίθεται κατώτερο όριο για τη διαθεσιμότητα. Αυτό ισοδυναμεί με τις προηγούμενες περιπτώσεις όπου το όριο της ελάχιστης διαθεσιμότητας τίθεται ίσο με 0%.
- Επιλογή της βέλτιστης διαθεσιμότητας. Σε αυτή την περίπτωση θα πρέπει να επιλέγονται για κάθε τύπο υπηρεσίας οι υπηρεσίες που προσφέρουν τη μεγαλύτερη διαθεσιμότητα.

5.4.3 Απαιτήσεις για το Χρόνο Εκτέλεσης

Αναφορικά με το χρόνο εκτέλεσης των εργασιών αναφέρεται στο χρόνο που οι πάροχοι δημοσιεύουν μέσω των συμβολαίων SLAs και μετριέται σε μονάδες χρόνου (για παράδειγμα 100 μονάδες χρόνου - ώρες). Οι κατηγορίες απαιτήσεων που έχουν σχέση με το χρόνο εκτέλεσης είναι οι εξής:

- Απαίτηση ανώτερου συνολικού χρόνου εκτέλεσης. Σε αυτή την περίπτωση ο χρήστης θέτει ένα ανώτατο όριο για το συνολικό χρόνο εκτέλεσης το οποίο δεν πρέπει να ξεπεραστεί. Αν αθροιστούν οι χρόνοι, που χρειάζονται για την εκτέλεση των υπηρεσιών που συμπεριλαμβάνονται σε μια ροή εργασίας, δε θα πρέπει να ξεπερνούν το ανώτατο αυτό όριο.
- Απαίτηση ανώτερου χρόνου ανά υπηρεσία. Σε αυτή την περίπτωση ο χρήστης καθορίζει ένα ανώτατο όριο χρόνου για κάθε τύπο υπηρεσίας. Ο χρόνος που απαιτείται για την εκτέλεση των υπηρεσιών δε θα πρέπει να ξεπερνάει το χρόνο που έχει οριστεί για κάθε τύπο υπηρεσίας.
- Αδιαφορία για το συνολικό χρόνο εκτέλεσης. Σε αυτή την περίπτωση ο χρήστης δεν ενδιαφέρεται για το συνολικό χρόνο εκτέλεσης της ροής εργασίας. Η τρίτη αυτή

περίπτωση ισοδυναμεί ουσιαστικά με τις προηγούμενες περιπτώσεις αλλά θέτοντας ως ανώτατο όριο χρόνου το άπειρο.

- Επιλογή του μικρότερου χρόνου εκτέλεσης. Σε αυτή την περίπτωση θα πρέπει να επιλέγονται για κάθε τύπο υπηρεσίας οι υπηρεσίες με το μικρότερο χρόνο εκτέλεσης ώστε ο συνολικός χρόνος εκτέλεσης που θα προκύψει για τη ροή εργασίας να είναι ο μικρότερος δυνατός.

5.4.4 Κατηγορίες Απαιτήσεων για όλες τις Παραμέτρους

Συνοπτικά οι κατηγορίες απαιτήσεων για όλες τις προαναφερθέντες παραμέτρους είναι οι ακόλουθες:

- Κατώφλια / Όρια. Σε αυτή την περίπτωση, ο χρήστης ορίζει κατώφλια στις τιμές των παραμέτρων. Αυτό σημαίνει ότι το συνολικό κόστος δεν πρέπει να υπερβεί έναν προκαθορισμένο προϋπολογισμό, ότι η διαθεσιμότητα πρέπει να είναι τουλάχιστον σε ένα συγκεκριμένο επίπεδο (εκφρασμένο με ποσοστιαία τιμή) και ότι πρέπει να τηρηθεί μια χρονική προθεσμία. Θα πρέπει επίσης να σημειωθεί ότι τα όρια μπορεί να ισχύουν ταυτόχρονα για μια ή περισσότερες παραμέτρους.
- Βελτιστοποίηση διαθεσιμότητας. Σε αυτή την περίπτωση, η επιλογή των υπηρεσιών γίνεται με την παραδοχή ότι ο κύριος παράγοντας επιλογής είναι η παράμετρος διαθεσιμότητάς τους και λόγω αυτού η αντιστοίχιση των ροών εργασίας γίνεται σε υπηρεσίες πλέγματος με τις υψηλότερες τιμές διαθεσιμότητας.
- Βελτιστοποίηση κόστους. Σε αυτή την περίπτωση, το κόστος των υπηρεσιών είναι ο κύριος παράγοντας κατά τη διαδικασία επιλογής. Το αποτέλεσμα της αντιστοίχισης είναι μια συγκεκριμένη ροή εργασίας (concrete workflow) που επιτυγχάνει το μικρότερο συνολικό κόστος.

- Βελτιστοποίηση χρόνου. Όμοια με τη διαθεσιμότητα και τη βελτιστοποίηση του κόστους, η επιλογή των υπηρεσιών γίνεται σε σχέση με το χρόνο εκτέλεσής τους προκειμένου να επιτευχθεί ο μικρότερος δυνατός συνολικός χρόνος εκτέλεσης για τη ροή εργασίας.
- Βέλτιστη λύση. Σε αυτή την περίπτωση, οι συντελεστές βαρύτητας για τις τρεις (3) παραμέτρους είναι ίσοι και οι υπηρεσίες που επιλέγονται προσφέρουν τη βέλτιστη τιμή διαθεσιμότητας και χρόνου εκτέλεσης για το αντίστοιχο κόστος τους.

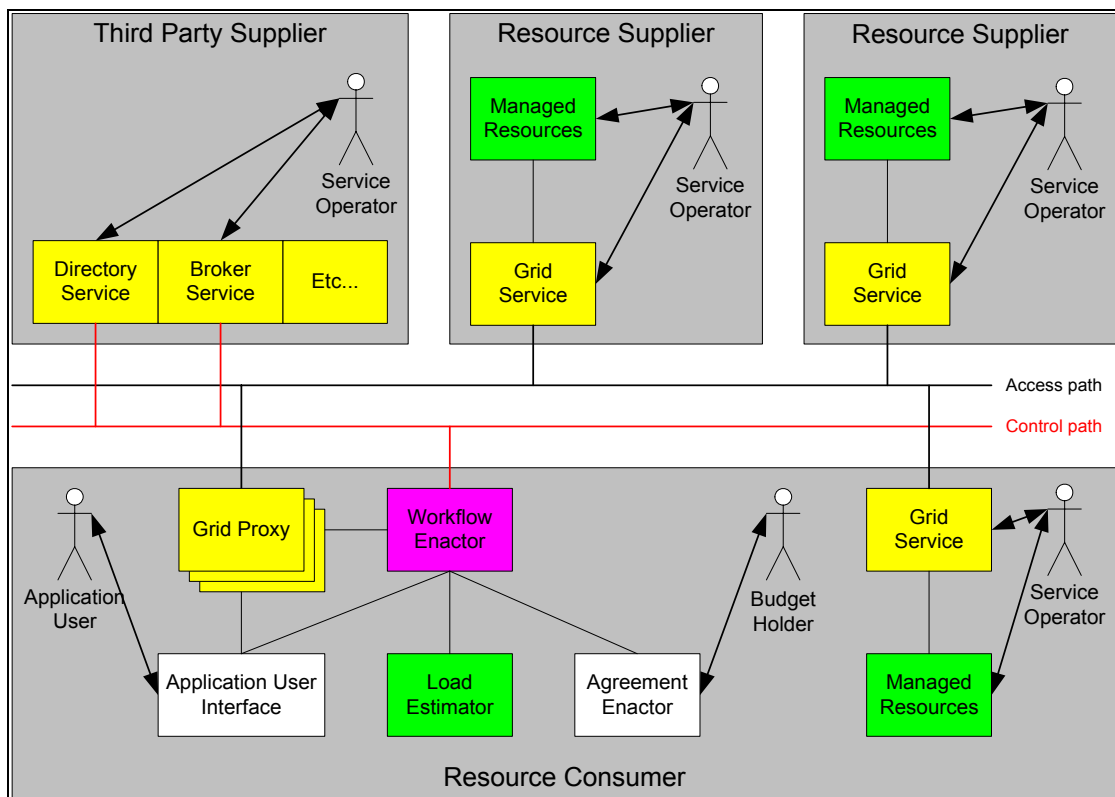
5.5 Μοντέλα Δρομολόγησης Εργασιών και Εκτίμησης Ποιότητας Υπηρεσίας

Στη συγκεκριμένη ενότητα παρουσιάζεται η έρευνα που ολοκληρώθηκε στις τεχνολογίες των αρχιτεκτονικών και υποδομών πλέγματος αναφορικά με τη δρομολόγηση εργασιών (με αλγόριθμους δρομολόγησης) και τη μελέτη της προσφερθείσας ποιότητας υπηρεσίας. Η προαναφερθείσα μελέτη βασίζεται στη μοντελοποίηση της αποδοτικότητας του συστήματος καθώς και στην εκτίμηση των απαιτήσεων διαφορετικών εργασιών σε υπολογιστικούς πόρους, με στόχο την εξαγωγή ποσοτικών δεικτών που αφορούν το βαθμό ικανοποίησης των απαιτήσεων των χρηστών μιας υποδομής, από τις παρεχόμενες υπηρεσίες.

5.5.1 Αρχιτεκτονική Υποδομής

Η επίτευξη της αίτησης για την ολοκλήρωση μιας εργασίας σε προδιαγεγραμμένο χρόνο απαιτεί από την αρχιτεκτονική του συστήματος κάποιες βασικές προϋποθέσεις, οι οποίες συσχετίζονται με τις παραμέτρους των πόρων και της εφαρμογής που επηρεάζουν το χρόνο εκτέλεσης. Η προαναφερθείσα αρχιτεκτονική περιλαμβάνει βασικές οντότητες, οι οποίες αναλύονται ακολούθως και παρουσιάζονται στο σχήμα της επόμενης σελίδας (Σχήμα 37):

- Προμηθευτές υπηρεσιών πλέγματος, που παρέχουν τις υπηρεσίες που εκτελούνται στους πόρους, έτσι ώστε να μπορούν να επιτύχουν αξιόπιστη ποιότητα υπηρεσίας.
- Καταναλωτής υπηρεσιών, που χρησιμοποιεί υπηρεσίες και ως εκ τούτου τους πόρους του πλέγματος
- Περιοχή τρίτων, όπου εκτελούνται υπηρεσίες όπως ένας κατάλογος αρχείων (όπου οι πάροχοι υπηρεσιών πλέγματος μπορούν να δημοσιεύσουν την ύπαρξη των υπηρεσιών τους) και ένας διαμεσολαβητής.
- Workflow enactor, ως τμήμα της υποδομής πλέγματος στην καταναλωτική περιοχή πόρων, ο οποίος αναλαμβάνει τη λήψη της περιγραφής της εργασίας που ο χρήστης θέλει να εκτελέσει, την συνδυάζει με μια ροή της δουλειάς επιχειρηματικής διαδικασίας, και εκτελεί το αποτέλεσμα. Ο workflow enactor έχει πρόσβαση στις απομακρυσμένες υπηρεσίες πλέγματος καλώντας τον αντίστοιχο Grid Proxy τοπικά.



Σχήμα 37: Αρχιτεκτονική περιγραφή υποδομής

Η σελίδα αυτή είναι σκόπιμα λευκή

Το βασικό μοντέλο της αρχιτεκτονικής για την επίτευξη της τήρησης του προδιαγεγραμμένου χρόνου εκτέλεσης απαιτεί την εκτίμηση του χρόνου που απαιτείται για την εκτέλεση μιας εργασίας σε κάποιο πόρο. Οι οντότητες που υλοποιούν το παραπάνω μοντέλο και οι οποίες επιτρέπουν την παροχή της απαιτούμενης ποιότητας υπηρεσίας όσον αφορά την τήρηση του προδιαγεγραμμένου χρόνου εκτέλεσης είναι οι εξής:

- Capacity Estimator, ο οποίος 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 και ο εκτιμητής πόρων λειτουργεί ως αντικειμενικό μέτρο σύγκρισης της απόδοσής τους.
- Workload Estimator, ο οποίος 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.
- Runtime Estimator. Οι παράμετροι φόρτου συνδυάζονται με τις παραμέτρους απόδοσης μέσω ενός μαθηματικού μοντέλου που καλύπτει τον υπολογισμό του χρόνου εκτέλεσης. Έτσι, είναι δυνατό για μία εργασία με γνωστές παραμέτρους φόρτου να υπολογιστεί ο χρόνος εκτέλεσής της σε δεδομένο πόρο.
- Scheduler, ο οποίος αναλαμβάνει την δρομολόγηση των εργασιών στους πόρους χρησιμοποιώντας την εκτίμηση για το χρόνο εκτέλεσης από το παραπάνω μοντέλο.

5.5.2 Παράμετροι που απαιτούνται για τη Δρομολόγηση

Οι παράμετροι που χρησιμοποιούνται στον υπολογισμό της χρονικής απόκρισης μιας υποδομής πλέγματος και απαιτούνται για τη δρομολόγηση αναφέρονται στις παραμέτρους

της απόδοσης, της διαθεσιμότητας των πόρων και της επικοινωνίας πελάτη – παρόχου.

Αναφορικά με την **απόδοση** διακρίνονται οι εξής περιπτώσεις:

- Απόδοση επεξεργαστή που μετράται σε MFLOPS (Million Floating point Operations Per Second) καθώς η μέτρηση με MIPS (Million Instructions Per Second) που αναφέρεται σε πόσες χαμηλού επιπέδου εντολές κώδικα μηχανής μπορεί να εκτελέσει ένας επεξεργαστής σε ένα δευτερόλεπτο δεν μπορεί να χρησιμοποιηθεί (κανένα τσιπ δεν χρησιμοποιεί ακριβώς το ίδιο είδος εντολών, μεθόδου εκτέλεσης, κ.λπ.). Επιπρόσθετα, στην περίπτωση που θέλουμε να εξετάσουμε την απόδοση του συστήματος, σημαντικό είναι να εξετάζουμε την υπάρχουσα ιεραρχία μνήμης (L1/L2 cache και κυρίως μνήμη RAM) και σε ποιο σημείο υπάρχει περιορισμός στην απόδοση. Η μέτρηση των MFLOPS πραγματοποιείται με μια διαδικασία που τρέχει χρησιμοποιώντας τις L1 και L2 cache. Επομένως, εάν μετράμε τα MFLOPS και επιπλέον το μέσο εύρος ζώνης γραφής/ανάγνωσης στην μνήμη RAM, μπορούμε να θεωρήσουμε ότι έχουμε λάβει υπόψη όλους τους παράγοντες που επηρεάζουν τον υπολογισμό.
- Απόδοση μονάδας σκληρού δίσκου, καθώς ακόμη και αν έχουμε πολύ γρήγορη μνήμη (θεωρητικά άπειρο εύρος ζώνης) παραμένει ένας ακόμα περιοριστικός παράγοντας στην μεταφορά δεδομένων προς τον επεξεργαστή, που είναι το εύρος ζώνης του σκληρού δίσκου. Καθώς είναι δύσκολο να προβλεφθεί η αλληλεπίδραση της εφαρμογής με το δίσκο θεωρούμε ως παράμετρο απόδοσης μόνο τον μέσο όρο του εύρους ζώνης γραφής/ανάγνωσης στο δίσκο.

Αναφορικά με τις παραμέτρους **διαθεσιμότητας** των πόρων (δείχνουν αν ο πόρος είναι διαθέσιμος να αναλάβει την εκτέλεση μιας στοιχειώδους εργασίας), διακρίνονται οι εξής περιπτώσεις:

- Ελάχιστη Ελεύθερη Μνήμη RAM (Minimum Free Memory), που είναι το ποσό της διαθέσιμης ελάχιστης ελεύθερης μνήμης που διαθέτει το σύστημα κατά την διάρκεια που πρέπει να γίνει ο υπολογισμός.
- Ελάχιστος Ελεύθερος χώρος στον σκληρό δίσκο (Minimum Free Disk Space), που αναφέρεται στον αποθηκευτικό χώρο που διαθέτει ο πόρος κατά τη διάρκεια που πρέπει να γίνει ο υπολογισμός.
- Χρονικό διάστημα αναμονής στην ουρά (Queue time interval). Σε μια υποδομή πλέγματος ο τοπικός διαχειριστής πόρων (Local Resource Manager) πρέπει να κρατάει για κάθε πόρο πληροφορίες σχετικά με την χρησιμοποίησή του από τις εφαρμογές. Εάν ο πόρος έχει επιλεγθεί για να τρέχει κάποια στοιχειώδη εργασία δεν είναι διαθέσιμος, παρά μόνο αν τελειώσει. Οι εργασίες που προορίζονται γι' αυτόν τον πόρο μπαίνουν σε μια ουρά αναμονής που το χρονικό μέγεθός της καθορίζεται από το άθροισμα των εκτιμώμενων χρόνων εκτέλεσης των στοιχειωδών εργασιών που την αποτελούν. Αυτό το άθροισμα είναι και το χρονικό διάστημα αναμονής στην ουρά για μια καινούρια στοιχειώδη εργασία.

Οι παράμετροι **επικοινωνίας** πελάτη-παρόχου εξετάζουν την ποιότητα σύνδεσης μεταξύ πελάτη και παρόχου η οποία είναι καθοριστική για την μεταφορά δεδομένων μεταξύ τους. Οι παράμετροι αυτοί διαχωρίζονται για εκπομπή και λήψη δεδομένων καθώς είναι πολύ πιθανό να υπάρχει διαφορετική ταχύτητα μεταξύ εκπομπής και λήψης. Οι παράμετροι επικοινωνίας περιγράφονται παρακάτω:

- Εύρος ζώνης εκπομπής (Send communication bandwidth), που αναφέρεται στο εύρος ζώνης εκπομπής της σύνδεσης του πελάτη με το διαδίκτυο (64K ISDN σύνδεση έχει εύρος ζώνης εκπομπής 64 Kbit/sec).

- Λανθάνων χρόνος εκπομπής (Send communication latency), που αναφέρεται στο χρόνο που απαιτείται από την εκπομπή ενός πακέτου από τον πελάτη μέχρι αυτό να φτάσει στον πάροχο.
- Εύρος ζώνης λήψης (Receive communication bandwidth) που αναφέρεται στο εύρος ζώνης λήψης της σύνδεσης του πελάτη με το διαδίκτυο.
- Λανθάνων χρόνος λήψης (Receive communication latency), που αναφέρεται στο χρόνο που απαιτείται από την εκπομπή ενός πακέτου από τον πάροχο μέχρι αυτό να φτάσει στον πελάτη.

5.5.3 Δρομολόγηση και παρεχόμενη Ποιότητα Υπηρεσίας

Η διαχείριση πόρων σε περιβάλλον πλέγματος έχει την ικανότητα να εγγυηθεί στις εμπορικές ή βιομηχανικές εφαρμογές μια εξατομικευμένη ποιότητα υπηρεσίας. Θεωρούμε την ποιότητα υπηρεσίας σχεδιάζοντας το πρόβλημα του οποίου ο στόχος είναι να καθορίσει πότε και σε ποιο πόρο μια δεδομένη στοιχειώδης εργασία πρέπει να εκτελεστεί, προκειμένου να ικανοποιηθούν οι ζητούμενες παράμετροι ποιότητας. Ο αλγόριθμος που μελετήθηκε για την δρομολόγηση χρησιμοποιεί μια Max-Min προσέγγιση δίκαιης διανομής για την παροχή δίκαιης πρόσβασης στους πόρους πλέγματος σε όλους τους χρήστες. Όταν δεν υπάρχει καμία έλλειψη πόρων, ο αλγόριθμός αναθέτει σε κάθε στοιχειώδη εργασία αρκετή υπολογιστική ισχύ για να ολοκληρωθεί μέσα στην προθεσμία της. Σε περίπτωση συμφόρησης, η βασική ιδέα είναι να μειώνει αρκετά τα ποσοστά CPU που ανατίθενται στις στοιχειώδεις εργασίες, έτσι ώστε το μερίδιο των πόρων που ο κάθε χρήστης παίρνει να είναι ανάλογο προς τη συμβολή των χρηστών στην υποδομή Grid ή εναλλακτικά στην τιμή που είναι πρόθυμοι να πληρώσουν. Ειδικότερα, όλες οι στοιχειώδεις εργασίες οι απαιτήσεις των οποίων είναι χαμηλότερες από το δίκαιο ποσοστό μεριδίου τους σε CPU εξυπηρετούνται στα απαιτημένα ποσοστά CPU τους. Εντούτοις, τα ποσοστά CPU στοιχειωδών εργασιών οι απαιτήσεις των

οποίων είναι μεγαλύτερες από το δίκαιο ποσοστό μεριδίου τους σε CPU μειώνονται για να κατανείμουν το συνολικό διαθέσιμο υπολογιστικό δυναμικό με έναν δίκαιο τρόπο.

5.5.4 Διαχείριση Πόρων και Δρομολόγηση

Οι αλγόριθμοι δρομολόγησης και τα σχέδια κατανομής πόρων εξαρτώνται από τον τύπο των επεξεργαστών και των στοιχειωδών εργασιών στο σύστημα, καθώς επίσης και από τους περιορισμούς δρομολόγησης που επιβάλλονται σε αυτά. Μια στοιχειώδης εργασία είναι «*non-preemptable*» εάν μόλις αρχίσει την εκτέλεσή της σε έναν επεξεργαστή, πρέπει να ολοκληρωθεί στον ίδιο επεξεργαστή. Επίσης, μια στοιχειώδης εργασία είναι «*non-interruptible*» εάν μόλις αρχίσει την εκτέλεσή της δεν πρέπει να διακοπεί από άλλες στοιχειώδεις εργασίες και να επαναλάβει την εκτέλεσή της αργότερα.

Στην ανάλυσή μας υποθέτουμε ότι οι στοιχειώδεις εργασίες είναι *non-preemptable* και *non-interruptible*, που είναι σαφώς η πιο περιοριστική περίπτωση, δεδομένου ότι οποιοσδήποτε αλγόριθμος δρομολόγησης που λειτουργεί με τις *non-preemptable* και *non-interruptible* στοιχειώδεις εργασίες θα λειτουργήσει επίσης όταν δεν απαιτείται ένας από αυτούς τους περιορισμούς.

5.5.5 Δίκαιη Δρομολόγηση

Η «δίκαιη» δρομολόγηση (*fair scheduling*) σημαίνει ότι ο αλγόριθμος δρομολόγησης δεν πρέπει να ευνοεί τις στοιχειώδεις εργασίες συγκεκριμένων χαρακτηριστικών (για παράδειγμα υψηλός φόρτος εργασίας) ενάντια σε άλλες. Προς αυτή την κατεύθυνση έχει προταθεί το σχήμα διανομής επεξεργαστή ή *Generalized Processor Sharing scheme* (GPS) [69], το οποίο εξομοιώνεται στην πράξη χρησιμοποιώντας το δίκαιο αλγόριθμο ζυγισμένης ουράς αναμονής ή *Weighted Fair Queuing* (WFQ) [54], [57].

Επιπρόσθετα, αρκετοί αλγόριθμοι δρομολόγησης έχουν αναφερθεί στην βιβλιογραφία. Ο πιο γνωστός αλγόριθμος σχεδιασμού είναι ο Earliest Deadline First (EDF) [58], [70] και [59] που υπαγορεύει ότι το σύστημα αναθέτει την υψηλότερη προτεραιότητα στη στοιχειώδη εργασία με την πιο επικείμενη προθεσμία. Άλλος αλγόριθμος δρομολόγησης είναι ο Slack Time, που αναφέρεται επίσης ως Least Laxity First (LLF) [70], [61] και [62], όπου οι στοιχειώδεις εργασίες επιλέγονται για την εκτέλεση με κατάταξη non-decreasing χαλαρού χρόνου, που καθορίζεται ως η διαφορά μεταξύ της σχετικής προθεσμίας της στοιχειώδους εργασίας και του εναπομένοντος υπολογιστικού χρόνου της. Η επιλογή της στοιχειώδους εργασίας μπορεί να γίνει επίσης χρησιμοποιώντας την ελάχιστη χρονική διάρκεια επεξεργασίας - Minimum Processing Time First approach (MPTF). Οι προαναφερθέντες αλγόριθμοι αποδίδουν ικανοποιητικά στην περίπτωση ενός ενιαίου επεξεργαστή, γιατί αν και καθορίζουν την κατάταξη στην οποία οι στοιχειώδεις εργασίες πρέπει να ανατεθούν προς εκτέλεση, δεν καθορίζουν τον επεξεργαστή στον οποίο οι επιλεγμένες στοιχειώδεις εργασίες ανατίθενται. Επίσης, δεν αντιμετωπίζουν τις στοιχειώδεις εργασίες κατά έναν δίκαιο τρόπο. Παραδείγματος χάριν, οι στοιχειώδεις εργασίες με σχετική επείγουσα ανάγκη (για τα σχήματα EDF και LLF) ή οι στοιχειώδεις εργασίες που έχουν μικρό φόρτο εργασίας μπορούν να ευνοηθούν ενάντια των υπολοίπων στοιχειωδών εργασιών. Ο αλγόριθμος που υλοποιήθηκε επιλύει τις προαναφερθείσες δυσκολίες καθώς οι στοιχειώδεις εργασίες που ζητούν εξυπηρέτηση είναι δρομολογημένες στους διαθέσιμους επεξεργαστές σύμφωνα με αυτό που καλούμε δίκαιους χρόνους ολοκλήρωσής τους (fair completion times). Ο δίκαιος χρόνος ολοκλήρωσης μιας στοιχειώδους εργασίας βρίσκεται πρώτα με τον υπολογισμό των δίκαια στοιχειωδών ποσοστών της (fair task rates) χρησιμοποιώντας έναν Max-Min δίκαιο αλγόριθμο διανομής.

Όπως έχει αναφερθεί παραπάνω, ένας αλγόριθμος δρομολόγησης πρέπει να απαντά σε δύο ερωτήσεις: στην κατάταξη των στοιχειωδών εργασιών προς επεξεργασία (το οποίο καλείται

πρόβλημα διάταξης ουράς αναμονής ή queue ordering problem) και στην απόφαση για τον επεξεργαστή στον οποίο πρέπει να ανατεθεί η εργασία (καλείται πρόβλημα ανάθεσης επεξεργαστών ή processor assignment). Για την επίλυση του πρώτου προβλήματος (**queue ordering problem**) στην δίκαιη δρομολόγηση, μελετήθηκαν αρκετοί κανόνες όπως περιγράφονται συνοπτικά στη συνέχεια:

- Απλή δίκαιη σειρά στοιχειωδών εργασιών (Simple Fair Task Order SFTO), όπου οι εργασίες ταξινομούνται στην ουρά με αύξουσα σειρά προτεραιότητας των μη ρυθμισμένων δίκαιων χρόνων ολοκλήρωσής τους. Οι μη ρυθμισμένες δίκαιες υπολογιστικές αναλογίες μιας εργασίας υπολογίζονται βασισμένες στις απαιτημένες υπολογιστικές αναλογίες τους και στην συνολική υπολογιστική δυνατότητα του πλέγματος. Τότε βρίσκεται ο μη ρυθμισμένος δίκαιος χρόνος ολοκλήρωσης της εργασίας και ταξινομούνται στην ουρά με αυτό το κριτήριο.
- Ρυθμιζόμενη δίκαιη σειρά στοιχειωδών εργασιών (Adjusted Fair Task Order AFTO), όπου οι στοιχειώδεις εργασίες διατάσσονται στην ουρά αναμονής κατά αύξουσα κατάταξη του ρυθμισμένου δίκαιου χρόνου ολοκλήρωσής τους. Το δίκαιο υπολογιστικό ποσοστό μιας στοιχειώδους εργασίας δεν είναι μια σταθερά, όπως υποτίθεται πριν, αλλά είναι στην πραγματικότητα μια συνάρτηση του χρόνου, η οποία αυξάνει όταν οι εργασίες τελειώνουν την εκτέλεσή τους και μειώνει όταν προκύπτουν νέες στοιχειώδεις εργασίες. Ο αλγόριθμος υπολογίζει αυτή τη χρονικά εξαρτημένη φύση των δίκαιων υπολογιστικών ποσοστών και επομένως τους ρυθμισμένους δίκαιους χρόνους ολοκλήρωσης.

Τα παραπάνω σχήματα (SFTO και AFTO) δίνουν την κατάταξη στην οποία οι στοιχειώδεις εργασίες πρέπει να εξεταστούν για την ανάθεση στους επεξεργαστές με έναν δίκαιο τρόπο. Παραμένει ακόμα να καθοριστούν οι επεξεργαστές στους οποίους οι στοιχειώδεις εργασίες ανατίθενται (processor assignment). Ο ένας τρόπος να ανατεθεί κάθε στοιχειώδης εργασία σε

έναν επεξεργαστή είναι με τη χρησιμοποίηση του πιο πρόωρου χρονικού σχεδίου ολοκλήρωσης (ECT). Στη συνέχεια παρέχουμε μια εναλλακτική προσέγγιση για τον προσδιορισμό του επεξεργαστή στον οποίο οι στοιχειώδεις εργασίες ανατίθενται. Η βασική ιδέα του αλγορίθμου είναι να συνδυαστούν αποτελεσματικά οι επεξεργαστές με την υπερχείλιση δυναμικού με τους επεξεργαστές με την περίσσεια ελεύθερου δυναμικού έτσι ώστε να επιτυγχάνεται μια καλύτερη εκμετάλλευση του γενικού δυναμικού επεξεργαστών. Βάσει αυτού, για όλους τους υπερχειλισμένους επεξεργαστές, πρέπει να διενεργηθεί μια μείωση των ποσοστών στοιχειωδών εργασιών. Η μείωση εκτελείται μέσω του δίκαιου τρόπου πάνω σε όλες τις στοιχειώδεις εργασίες που ανατίθενται στον υπερχειλισμένο επεξεργαστή. Σε περίπτωση που όλοι οι χρήστες που υποβάλλουν τις στοιχειώδεις εργασίες, συμβάλλουν εξίσου στην υποδομή πλέγματος, η υπερχείλιση επεξεργαστών μοιράζεται εξίσου μεταξύ όλων των στοιχειωδών εργασιών του επεξεργαστή, το οποίο οδηγεί σε μια δίκαιη πολιτική υπερχείλισης επεξεργαστών. Για την εφαρμογή του προαναφερθέντος σχήματος, απαιτείται μια αρχική ρύθμιση στοιχειωδών εργασιών. Αυτό εξετάζεται σε αυτό το τμήμα, χρησιμοποιώντας μια μέθοδο παρόμοια με αυτή που χρησιμοποιείται στο πρόβλημα bin packing, λόγω του γεγονότος ότι έχει μελετηθεί ευρέως στην βιβλιογραφία και μπορούν να βρεθούν τα θεωρητικά όρια όσον αφορά τη βέλτιστη λύση. Ειδικότερα, ο αλγόριθμος τακτοποιεί τις στοιχειώδεις εργασίες όσον αφορά το ποσοστό τους. Η στοιχειώδης εργασία του μέγιστου ποσοστού δρομολογείται αρχικά ακολουθούμενη από τη δεύτερη μεγαλύτερη στοιχειώδη εργασία και ούτω καθεξής. Οι στοιχειώδεις εργασίες ανατίθενται σε έναν επεξεργαστή έτσι ώστε α) να μπορούν να δρομολογηθούν εφικτά και β) να επιτυγχάνεται η βέλτιστη τακτοποίηση του δυναμικού επεξεργαστών. Η πρώτη συνθήκη σημαίνει ότι το διαθέσιμο δυναμικό επεξεργαστών είναι αρκετό για να δρομολογηθεί η στοιχειώδης εργασία. Η δεύτερη συνθήκη δείχνει ότι το υπόλοιπο χρησιμοποιούμενο δυναμικό, μετά από την ρύθμιση των στοιχειωδών εργασιών, πρέπει να είναι όσο το δυνατόν

χαμηλότερο. Σε περίπτωση που μια επιλεγμένη στοιχειώδης εργασία δεν μπορεί να δρομολογηθεί εφικτά σε οποιοδήποτε επεξεργαστή, η στοιχειώδης εργασία ανατίθεται στον επεξεργαστή στον οποίο αντιμετωπίζεται η ελάχιστη υπερχείλιση.

Το κεντρικό στοιχείο οποιασδήποτε αρχιτεκτονικής πλέγματος είναι το διαχειριστικό εργαλείο των πόρων που χρησιμοποιείται για να δρομολογήσει τις στοιχειώδεις εργασίες στους διαθέσιμους πόρους, με χρήση παραμέτρων απόδοσης, διαθεσιμότητας και επικοινωνίας πελάτη - παρόχου. Βάσει αυτών, η επιτυχής δρομολόγηση θεωρείται βασικό σημείο για την επιτυχία των υπολογιστικών πλεγμάτων, και την ικανότητα αυτών να υποστηρίζουν εμπορικές εφαρμογές εξασφαλίζοντας την παρεχόμενη ποιότητα υπηρεσίας.

Η σελίδα αυτή είναι σκόπιμα λευκή

6

Απεικόνιση Δυναμικών Ροών

Εργασίας και Επιλογή

Υπηρεσιών

Στο κεφάλαιο αυτό περιγράφεται το μοντέλο και ο μηχανισμός απεικόνισης δυναμικών ροών εργασίας σε υπηρεσίες πλέγματος με στόχο την παροχή ποιότητας υπηρεσιών. Γίνεται προσδιορισμός του προβλήματος και προτείνεται ένας αλγόριθμος για την προαναφερθείσα απεικόνιση ενώ παράλληλα περιγράφεται ένα σχήμα δημιουργίας περισσοτέρων ροών εργασίας προς όφελος του τελικού χρήστη. Επίσης, περιγράφεται η ενσωμάτωση του συγκεκριμένου μηχανισμού στην αρχιτεκτονική που παρουσιάστηκε στο τέταρτο κεφάλαιο (βλ. ενότητα 4.4) της παρούσης διατριβής.

6.1 Παρούσα Κατάσταση – Σχετικές Εργασίες

Με στόχο την εξασφάλιση παροχής ποιότητας υπηρεσίας με τη χρήση μηχανισμών απεικόνισης ροών εργασίας, υπάρχουν διάφορες προσεγγίσεις για τον τρόπο διαχείρισης των πληροφοριών σχετικά με την προσφερόμενη ποιότητα. Επίσης, υπάρχουν συμβατικές μέθοδοι επιλογής και χρονοπρογραμματισμού. Στις βιβλιογραφικές αναφορές [22] και [23]

περιγράφονται οι αλγόριθμοι OLB, UDA, Suffrage, GSA, Tabu κλπ με βασικό στόχο το μικρότερο δυνατό άνοιγμα, όπου λαμβάνεται υπόψη μόνο η απόδοση του συστήματος αλλά δεν συνυπολογίζεται η απαίτηση του χρήστη για ποιότητα υπηρεσιών.

Στις αναφορές [24] και [25] παρουσιάζονται τρεις αλγόριθμοι όπου χρησιμοποιούνται δυο παράμετροι – κόστος και χρόνος εκτέλεσης – προκειμένου να εκφραστεί η διάσταση της ποιότητας υπηρεσιών. Αυτές οι παράμετροι χρησιμοποιούνται στην υλοποίηση ενός μηχανισμού επιλογής, που αναφέρεται σε εφαρμογές που αποτελούνται από παραμετρικές διαδικασίες που είναι ανεξάρτητες μεταξύ τους αναφορικά με τη σειρά εκτέλεσής τους. Ο πρώτος αλγόριθμος ελαχιστοποιεί το κόστος, ο δεύτερος ελαχιστοποιεί το κόστος σε σχέση με το χρόνο και ο τρίτος είναι γενικός και δεν στοχεύει στη βελτιστοποίηση κάποιας παραμέτρου αλλά εκχωρεί στους "φθηνότερους" κόμβους έναν αριθμό εργασιών που είναι αντιστρόφως ανάλογος ως προς το χρόνο που χρειάζεται κάθε κόμβος για την ολοκλήρωση μιας εργασίας.

Παρόμοια, στην αναφορά [26] περιγράφεται ένας αλγόριθμος χρονοπρογραμματισμού που λαμβάνει υπόψη δυο (2) κύριες παραμέτρους (κόστος και χρόνο). Περιλαμβάνονται τέσσερις προσαρμόσιμοι αλγόριθμοι χρονοπρογραμματισμού με περιορισμούς ως προς το χρόνο εκτέλεσης και το κόστος. Αυτοί οι αλγόριθμοι χρησιμοποιούνται για τη βελτιστοποίηση του κόστους – στα πλαίσια των περιορισμών χρόνου και προϋπολογισμού, για τη βελτιστοποίηση του χρόνου – στα πλαίσια των περιορισμών χρόνου και προϋπολογισμού, για συντηρητική βελτιστοποίηση του χρόνου – στα πλαίσια των περιορισμών χρόνου και προϋπολογισμού και για τη βελτιστοποίηση κόστους και χρόνου – στα πλαίσια των περιορισμών χρόνου και προϋπολογισμού.

Επίσης, στην αναφορά [27] παρουσιάζεται ένας τρόπος διαχείρισης της ποιότητας υπηρεσίας μιας ροής εργασίας και μέθοδοι πρόβλεψης, ανάλυσης και παρακολούθησης της ποιότητας. Γίνεται εκτίμηση της ποιότητας για συγκεκριμένα μεγέθη - χρόνος, κόστος και αξιοπιστία -

με τη χρήση δυο μεθόδων: ανάλυσης και προσομοίωσης. Σε αυτή την περίπτωση η διαχείριση των παραμέτρων γίνεται μεμονωμένα, όπως στις αναφορές [24] και [25], και όχι με συνδυασμένο τρόπο ενώ η συνολική αξιολόγηση προκύπτει από τις επιμέρους εργασίες, γεγονός που καθιστά την παροχή ποιότητας υπηρεσιών για όλη τη ροή εργασίας ανέφικτη καθώς η αξιολόγηση θα έπρεπε να πραγματοποιείται με αντίστροφο τρόπο.

Στις αναφορές [28] και [29] προτείνονται ορισμένοι αλγόριθμοι χρονοπρογραμματισμού βελτιστοποίησης με τον περιορισμό του χρόνου και του κόστους στο μοντέλο Nimrod-G. Το Nimrod-G είναι βασισμένο στους όρους της αγοράς σύστημα χρονοπρογραμματισμού και διαχείρισης πόρων για περιβάλλοντα πλέγματος [29]. Επιπλέον ενσωματώνει το πρόγραμμα Libra το οποίο χρησιμοποιεί έναν χρονοπρογραμματιστή για τον υπολογισμό συστάδων (cluster) [92]. Υποστηρίζει περιορισμούς χρονοπροθεσμίας και προϋπολογισμού που ορίζονται από το χρήστη για τη ρύθμιση της παροχής και ζήτησης πόρων στο πλέγμα με αξιοποίηση των υπηρεσιών ανταλλαγής πόρων GRACE (Grid Architecture for Computational Economy) [100], [101]. Το Nimrod-G [30] επιτρέπει στους χρήστες να παραχωρούν και να συναθροίζουν υπηρεσίες πόρων ανάλογα με τη διαθεσιμότητα, τη δυναμικότητα, την απόδοση και το κόστος τους καθώς και τις απαιτήσεις του χρήστη για ποιότητα υπηρεσίας. Το DAGMan [32] είναι ένα σύνολο βιβλιοθηκών C που δίνει στο χρήστη δυνατότητα χρονοπρογραμματισμού υπηρεσιών με βάση σχέσεις εξάρτησης. Το DAGMan είναι μέρος του έργου Condor και επεκτείνει το *Condor Job Scheduler* (μηχανισμός χρονοπρογραμματισμού εργασιών Condor) προκειμένου να διαχειρίζεται εξαρτήσεις μεταξύ εργασιών. Όπως υποδηλώνεται από το όνομά του, το DAGMan απεικονίζει ένα σύνολο εξαρτήσεων εργασιών ως κατευθυντικό ακυκλικό γράφημα. Χρησιμοποιεί το DAG ως δομή δεδομένων για την απεικόνιση των εξαρτήσεων των εργασιών. Άλλες ενδιαφέρουσες εργασίες με θέμα τους αλγορίθμους επιλογής και χρονοπρογραμματισμού παρουσιάζονται στις αναφορές [33] και [34], όπου στην πρώτη οι

συγγραφείς παρουσιάζουν έναν ευρετικό αλγόριθμο χρονοπρογραμματισμού εργασιών με περιορισμούς στην ποιότητα υπηρεσίας, ενώ η δεύτερη διαχειρίζεται το πρόβλημα χρονοπρογραμματισμού κατανεμημένων εργασιών σε πλέγματα με χρήση πολλών ταυτόχρονων αιτημάτων. Παράλληλα στη βιβλιογραφία [89], [90] και [91] συναντώνται θέματα που αφορούν στην πρόβλεψη του χρόνου εκτέλεσης εργασιών και χρησιμοποιούνται για τη δρομολόγηση και την παροχή ποιότητας υπηρεσίας. Τέλος, έχουν υλοποιηθεί μηχανισμοί που ενσωματώνουν πολιτικές χρονοδρομολόγησης με παράλληλη εκτίμηση κόστους όπως αναφέρεται στις [83], [84] και [86].

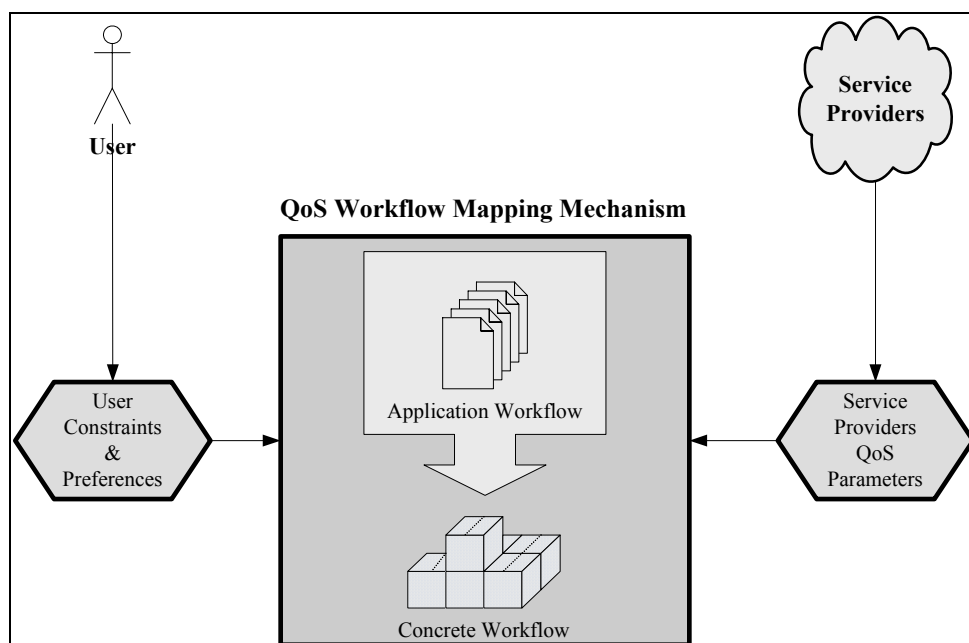
Η διαφορά μεταξύ των συστημάτων που παρουσιάζονται σε αυτή την ενότητα και του μηχανισμού που αναπτύχθηκε και παρουσιάζεται στη συνέχεια έγκειται στο ότι τα συστήματα που παρουσιάστηκαν παραπάνω αναφέρονται στην επιλογή υπηρεσιών και κόμβων με βάση τις παραμέτρους της ποιότητας υπηρεσίας αντιμετωπίζοντας μόνο τις ειδικές περιπτώσεις ελαχιστοποίησης μιας από τις παραμέτρους ενώ παράλληλα όλες οι παράμετροι έχουν τον ίδιο συντελεστή βαρύτητας. Επιπρόσθετα αν και οι δυνατότητες πολλών συστημάτων είναι αξιόλογες, πολλά από αυτά δεν είναι συμβατά με την αρχιτεκτονική OGSA (βλ. ενότητα 4.3). Στο μηχανισμό που παρουσιάζεται στα πλαίσια της παρούσας διδακτορικής διατριβής, οι παράμετροι που εκφράζουν την ποιότητα υπηρεσίας αντιμετωπίζονται και με συνδυασμένο τρόπο ενώ παράλληλα εισάγεται η περίπτωση όπου ο χρήστης ορίζει τις προτιμήσεις του και επομένως μια από τις παραμέτρους μπορεί να παίξει σημαντικότερο ρόλο στη διαδικασία επιλογής (της εκχωρείται συντελεστής βαρύτητας).

6.2 Καινοτόμος Μηχανισμός Απεικόνισης Δυναμικών Ροών

Εργασίας και Επιλογής Υπηρεσιών Πλέγματος

Όπως αναφέρθηκε παραπάνω, ο σκοπός των μηχανισμών απεικόνισης δυναμικών ροών εργασίας και επιλογής υπηρεσιών πλέγματος για διαχείριση της προσφερθείσας ποιότητας

υπηρεσίας είναι ο προσδιορισμός και η περιγραφή των διαδικασιών που πρέπει να ολοκληρωθούν προκειμένου να οριστεί η συγκεκριμένη ροή εργασίας (concrete workflow) με δεδομένα τη ροή εργασίας μιας εφαρμογής (application workflow) καθώς και τις παραμέτρους και τις προτιμήσεις της ποιότητας υπηρεσίας. Κάθε ροή εργασίας περιλαμβάνει τύπους υπηρεσιών που μπορούν να εκτελεστούν από ένα σύνολο υπηρεσιών (υποψήφιος), οι οποίες συνοδεύονται από πληροφορίες σχετικά με την παρεχόμενη ποιότητα υπηρεσίας εκφρασμένες με συγκεκριμένες παραμέτρους (όπως αυτές περιγράφηκαν στην ενότητα 5.3). Ο μηχανισμός απεικόνισης της ροής εργασίας (όπως αυτός παρουσιάζεται και στις αναφορές [42] και [43]) επιτρέπει την επιλογή των υποστάσεων υπηρεσιών (υποψηφίων) για κάθε τύπο υπηρεσίας με βάση τη ροή εργασίας της εφαρμογής (application workflow), τους περιορισμούς και τις προτιμήσεις του χρήστη καθώς και τις παραμέτρους QoS για κάθε υπηρεσία. Τα δεδομένα εισόδου του προαναφερθέντος μηχανισμού παρουσιάζονται στο σχήμα που ακολουθεί (Σχήμα 38) και αναλύονται στη συνέχεια:



Σχήμα 38: Συνοπτική παρουσίαση μηχανισμού απεικόνισης ροών εργασίας

Η σελίδα αυτή είναι σκόπιμα λευκή

- Ροή εργασίας εφαρμογής με πληροφορίες Ταξινόμησης, που υποδηλώνουν ποιος από τους τύπους υπηρεσιών παίζει σημαντικότερο ρόλο στη ροή εργασίας. Αυτός ο παράγοντας πρέπει να συμπεριληφθεί στη ροή εργασίας της εφαρμογής η οποία παρέχεται από τους σχεδιαστές της εφαρμογής (για παράδειγμα ποια υπηρεσία είναι υψηλότερης σημασίας: μετακίνηση περισσότερων δεδομένων ή χρόνος εκτέλεσης) καθώς υποδηλώνει τη σειρά εκτέλεσης των τύπων υπηρεσιών. Το τελικό αποτέλεσμα αυτού του βήματος είναι η κατασκευή της αφηρημένης ροής εργασίας (abstract workflow), με τις διαθέσιμες υπηρεσίες ανά τύπο υπηρεσίας.
- Περιορισμοί και προτιμήσεις Ποιότητας Υπηρεσιών που ορίζονται από το χρήστη. Ένα σύνολο περιορισμών (για παράδειγμα κατώφλια / όρια σε συγκεκριμένες τιμές των παραμέτρων) μαζί με πιθανές προτιμήσεις του χρήστη, που αποδίδουν ένα συντελεστή βαρύτητας στις παραμέτρους ποιότητας και αναλύονται μέσα στο μηχανισμό απεικόνισης της ροής εργασίας.
- Πληροφορίες Ποιότητας Υπηρεσιών για τις υπηρεσίες πλέγματος. Κάθε υπηρεσία ενός τύπου υπηρεσιών συνοδεύεται από πληροφορίες ποιότητας (παραμέτρους) που θα χρησιμοποιηθούν κατά τη διαδικασία επιλογής, όπως αυτές δημοσιεύονται από τους παρόχους των υπηρεσιών.

6.3 Προσδιορισμός του Προβλήματος και Λύση

Στις παραγράφους που ακολουθούν γίνεται προσδιορισμός του προβλήματος επιλογής υπηρεσιών και απεικόνισης ροών εργασίας σε υπηρεσίες που προσφέρονται από περιβάλλοντα πλέγματος καθώς και παρουσίαση της λύσης αυτού.

6.3.1 Αντικατάσταση Διαθεσιμότητας, Χρόνου και Κόστους από ένα Κοινό Μέγεθος

Το πρόβλημα που καλείται να επιλύσει ένας μηχανισμός απεικόνισης ροών εργασίας και επιλογής υπηρεσιών παρουσιάζει πολλές ομοιότητες με την κατηγορία προβλημάτων εύρεσης του συντομότερου μονοπατιού (shortest path problem). Σε αυτή την περίπτωση κρίνεται χρήσιμη η αντικατάσταση του κόστους χρήσης ενός πόρου, του χρόνου εκτέλεσης και της διαθεσιμότητας που προκύπτει για αυτόν από ένα κοινό μέγεθος. Λόγω αυτού απαιτείται ο υπολογισμός του λόγου του κόστους και χρόνου προς τη διαθεσιμότητα που ουσιαστικά αντιστοιχεί στο κόστος και το χρόνο ανά μονάδα διαθεσιμότητας για κάθε υποψήφια υπηρεσία πλέγματος. Ο λόγος αυτός θα αποτελεί οδηγό για την επιλογή μεταξύ δύο υπηρεσιών ίδιου τύπου. Προφανώς ανάμεσα σε δύο στιγμιότυπα υπηρεσίας θα πρέπει να επιλέγεται αυτό με το μικρότερο κόστος και χρόνο εκτέλεσης ανά μονάδα διαθεσιμότητας (εφόσον βέβαια δεν υπάρχουν συγκεκριμένες προτιμήσεις για τις παραπάνω παραμέτρους όπως περιγράφηκε στην ενότητα 6.2 και επομένως να επιτυγχάνεται η μεγαλύτερη δυνατή διαθεσιμότητα με το μικρότερο δυνατό κόστος και χρόνο εκτέλεσης ενώ παράλληλα θα πρέπει να τηρούνται τα όρια του χρήστη για τις παραμέτρους.

Βασικό κριτήριο θα πρέπει να είναι η διαθεσιμότητα καθώς εάν μια υπηρεσία δεν ικανοποιεί το όριο για τη διαθεσιμότητα, τότε θα επιλέγεται κάποιο άλλη ακόμη και αν ο λόγος της νέας επιλογής είναι μεγαλύτερος από αυτόν την αρχικής. Αντίστοιχα, εάν η επιλογή μιας υπηρεσίας οδηγεί σε παραβίαση του περιορισμού συνολικού κόστους, θα πρέπει να επιλεγεί

κάποια άλλη ώστε ο περιορισμός να ικανοποιείται ακόμη και αν ο λόγος κόστος και χρόνος / διαθεσιμότητα την καθιστά λιγότερο συμφέρουσα. Αναφορικά με την περίπτωση του χρόνου, ο συνολικός χρόνος εκτέλεσης μιας ροής εργασίας θα αντιμετωπίζεται συνολικά με την εφαρμογή ενός σχήματος κατανομής και παραλληλοποίησης των εργασιών ώστε να επιτυγχάνεται μείωση αυτού και εκπλήρωση τυχόν ορίου που μπορεί να έχει τεθεί από το χρήστη.

6.3.2 Ανάγκη Τροποποίησης του Λόγου Κόστους και Χρόνου / Διαθεσιμότητα

Μέχρι στιγμής έχει γίνει αναλυτική αναφορά στο κόστος, στο χρόνο και στη διαθεσιμότητα ως παραμέτρους για την επιλογή των υπηρεσιών. Εκτός όμως από τις παραπάνω παραμέτρους υπάρχει και μια ακόμη που αφορά στο εάν ενδιαφέρει το χρήστη περισσότερο το κόστος, ο χρόνος εκτέλεσης ή η διαθεσιμότητα. Για παράδειγμα, μπορεί να είναι επιθυμητό να δοθεί περισσότερο βάρος στην εύρεση μιας λύση χαμηλού κόστους ακόμη και αν υπάρχουν αναλογικά σημαντικές αρνητικές επιπτώσεις στη διαθεσιμότητα. Δηλαδή είναι δυνατόν ο χρήστης να προτιμά να έχει μια οικονομικότερη λύση με μικρότερη διαθεσιμότητα, αντί μιας λιγότερο οικονομικής λύσης για την επιλογή και χρήση μιας υπηρεσίας ακόμη και αν αυτή προσφέρει τέτοια διαθεσιμότητα, ώστε ο λόγος κόστους και χρόνου / διαθεσιμότητα να είναι μικρότερος.

Βάσει των παραπάνω είναι προφανές ότι ο απλός λόγος κόστους και χρόνου προς τη διαθεσιμότητα είναι ανεπαρκής για να εκφράσει την επιπλέον παράμετρο ως προτίμηση του χρήστη. Επομένως θα πρέπει να τροποποιηθεί ώστε να είναι δυνατόν να ικανοποιούνται και αυτή η επιπλέον παράμετρος του ενδιαφέροντος για το κόστος, το χρόνο εκτέλεσης και τη διαθεσιμότητα.

6.3.3 Τροποποιημένος Λόγος Κόστους και Χρόνου / Διαθεσιμότητα

Σε συνέχεια των προηγούμενων, ο τροποποιημένος λόγος κόστους και χρόνου / διαθεσιμότητα προκύπτει από την πρόσθεση των παραμέτρων για το κόστος, το χρόνο εκτέλεσης και τη διαθεσιμότητα, οι οποίες θα λειτουργούν ως βάρη για τα αντίστοιχα μεγέθη. Τα βάρη αυτά θα μεταβάλλονται ανάλογα με το εάν ο χρήστης εκφράσει ενδιαφέρον για το κόστος, το χρόνος ή τη διαθεσιμότητα. Επίσης θα πρέπει να μεταβάλλονται και σύμφωνα με το ύψος του κόστους, του χρόνου και της διαθεσιμότητας. Σκοπός είναι με τη μεταβολή των παραμέτρων να προσαρμόζεται ο λόγος έτσι ώστε να αντανakλά τις πραγματικές προτιμήσεις. Βάσει αυτών, ο λόγος γίνεται:

$$\frac{(f_{Cost} * Cost) * (f_{Time} * Time)}{f_{Availability} * Availability} \quad (7)$$

Το βάρος f_{Cost} είναι αυτό που μεταβάλλει το κόστος, το βάρος f_{Time} αυτό που μεταβάλλει το χρόνο και το βάρος $f_{Availability}$ αυτό που μεταβάλλει τη διαθεσιμότητα. Με το ανωτέρω εννοούμε ότι τελικά το κόστος, ο χρόνος ή αντίστοιχα η διαθεσιμότητα, που θα συμμετέχει στον αρχικό τύπο, θα παρουσιάζεται τροποποιημένο σε σχέση με το αυθεντικό. Δηλαδή το βάρος f_{Cost} θα ενισχύει τη συμβολή του κόστους στη διαμόρφωση του λόγου, με αποτέλεσμα στο τελικό κλάσμα να εμφανίζεται ένα πλασματικό κόστος. Αντίστοιχα και τα βάρη f_{Time} και $f_{Availability}$ θα ενισχύουν τη συμβολή του χρόνου εκτέλεσης και της διαθεσιμότητας στη διαμόρφωση του λόγου, με αποτέλεσμα στο τελικό κλάσμα να εμφανίζεται ένας πλασματικός-τροποποιημένος χρόνος ή διαθεσιμότητα. Το μέγεθος της ενίσχυσης και για τα τρία (3) μεγέθη θα εξαρτάται από δύο παράγοντες:

- Για ποιο από τα μεγέθη έχει εκφράσει μεγαλύτερο ενδιαφέρον ο χρήστης. Εάν για παράδειγμα έχει εκφράσει για το κόστος θα πρέπει το βάρος f_{Cost} να τείνει να ενισχύσει το κόστος περισσότερο απ' ό,τι τα βάρη f_{Time} και $f_{Availability}$ θα τείνουν να ενισχύσουν το χρόνο και τη διαθεσιμότητα αντίστοιχα.

- Την τιμή του κάθε μεγέθους. Αυτό σημαίνει ότι για παράδειγμα εάν υπάρχουν δύο προσφορές, από τις οποίες η μία έχει μεγαλύτερο κόστος από την άλλη, τότε η ενίσχυση που θα εισάγει το f_{Cost} θα είναι διαφορετική για το κάθε κόστος. Ουσιαστικά όσο μεγαλύτερο είναι το κόστος τόσο μεγαλύτερη θα πρέπει να είναι και η ενίσχυση. Αντίστοιχα το ίδιο ισχύει και στην περίπτωση του χρόνου εκτέλεσης και της διαθεσιμότητας.

6.3.4 Η Συνάρτηση Μετατροπής

Με βάση όσα αναφέρθηκαν στην προηγούμενη παράγραφο είναι προφανές ότι τα f_{Cost} , f_{Time} και $f_{\text{Availability}}$ δεν έχουν σταθερή τιμή αλλά αποτελούν συναρτήσεις, για τις οποίες πρέπει να καθοριστούν οι μεταβλητές και τα υπόλοιπα χαρακτηριστικά τους (για παράδειγμα αν θα είναι αύξουσες, ποιοι θα είναι οι σταθεροί όροι που θα περιλαμβάνουν κ.α.). Θα πρέπει στο σημείο αυτό να σημειωθεί ότι δεν υπάρχει μια συγκεκριμένη συνάρτηση που να μοντελοποιεί πιστά τον τρόπο με τον οποίο ο χρήστης αντιμετωπίζει το κόστος, το χρόνο ή τη διαθεσιμότητα, καθώς εμπλέκεται το υποκειμενικό στοιχείο. Παρόλα αυτά είναι δυνατόν όμως να δημιουργηθούν συναρτήσεις που δίνουν αρκετά ικανοποιητικά αποτελέσματα, όσον αφορά το πώς η προτίμηση στα προαναφερθέντα μεγέθη επηρεάζει τελικά το λόγο τους.

Στη συνέχεια παρουσιάζεται η γενική μορφή της συνάρτησης με την οποία προκύπτουν τα f_{Cost} , f_{Time} και $f_{\text{Availability}}$, τα χαρακτηριστικά της οποίας είναι:

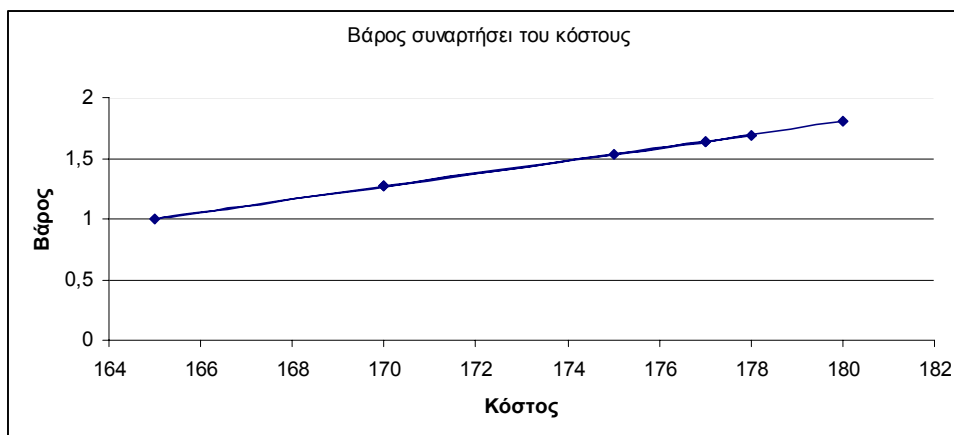
- Η συνάρτηση f_{Cost} έχει ως μεταβλητές το κόστος της προσφοράς και το CostSlope που σχετίζεται με το ενδιαφέρον του χρήστη για το κόστος. Αντίστοιχα οι συναρτήσεις f_{Time} και $f_{\text{Availability}}$ έχουν ως μεταβλητές το χρόνο εκτέλεσης και τη διαθεσιμότητα της προσφοράς και τα μεγέθη TimeSlope και AvailabilitySlope που θα σχετίζεται με το ενδιαφέρον για το χρόνο και τη διαθεσιμότητα αντίστοιχα.

- Πεδίο ορισμού για το κόστος, το χρόνο και τη διαθεσιμότητα. Λόγω της φύσης των υπηρεσιών πλέγματος το κόστος είναι δυνατόν να είναι μόνο θετικό καθώς δεν υπάρχουν προσφορές υπηρεσίας με αρνητικό κόστος. Όσον αφορά το χρόνο και τη διαθεσιμότητα, επίσης όπως και το κόστος, είναι δυνατόν να έχουν μόνο θετικές τιμές, με τη διαθεσιμότητα να λαμβάνει τιμές που έχουν ως άνω όριο το 100, καθώς έχει ήδη αναφερθεί ότι η διαθεσιμότητα υπολογίζεται σε μονάδες επί τοις εκατό.
- Οι συναρτήσεις f_{Cost} , f_{Time} και $f_{\text{Availability}}$ θα πρέπει να είναι αύξουσες συναρτήσεις. Αυτό στην περίπτωση του κόστους σημαίνει ότι μεγάλες τιμές κόστους θα τείνουν να γίνουν πολύ μεγαλύτερες από όσο τείνουν οι μικρές τιμές κόστους. Το αντίστοιχο ισχύει και για το χρόνο και τη διαθεσιμότητα. Έστω ότι έχουμε μια γραμμική συνάρτηση, τότε ένα παράδειγμα γραφικής παράστασης, που αναπαριστά το βάρος που αντιστοιχεί σε κάθε κόστος υπηρεσίας εμφανίζεται στην επόμενη σελίδα (Σχήμα 39).

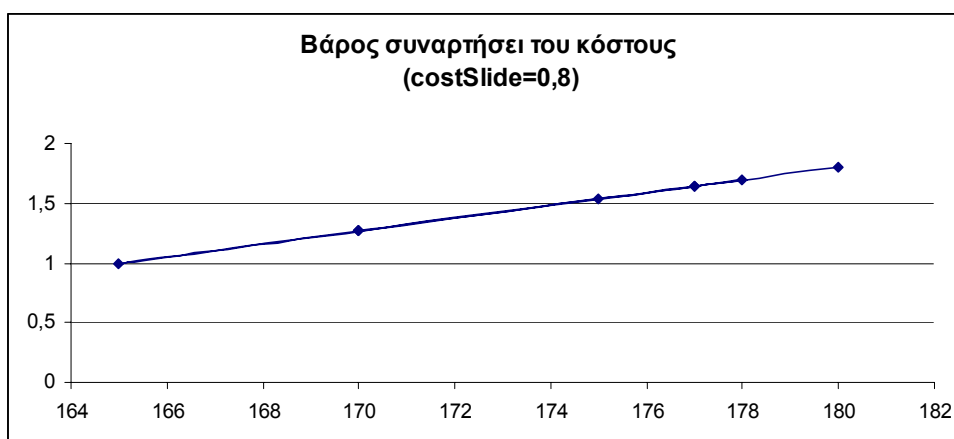
Το διάγραμμα αφορά τα βάρη που αντιστοιχούν στα κόστη των προσφορών για μια υπηρεσία και είναι ενδεικτικό, όπου φαίνεται ότι σε μεγαλύτερο κόστος αντιστοιχεί και μεγαλύτερο βάρος. Αυτό σημαίνει ότι όσο μεγαλύτερο είναι το κόστος, τόσο ακόμη μεγαλύτερο θα είναι το πλασματικό κόστος που αντιλαμβάνεται ο χρήστης. Αντίστοιχα ισχύουν και για το χρόνο και τη διαθεσιμότητα.

Οι αυξητικές τάσεις της συνάρτησης θα πρέπει να καθορίζονται από το ενδιαφέρον για το κόστος (CostSlope), τη διαθεσιμότητα (AvailabilitySlope) και το χρόνο (TimeSlope). Αυτό σημαίνει ότι η καθεμία από τις παραπάνω παραμέτρους θα πρέπει να διατηρείται αν γίνει παραγωγή της αντίστοιχης συνάρτησης. Επίσης όσο μεγαλύτερη τιμή έχει, τόσο μεγαλύτερη θα πρέπει να είναι και η αυξητική τάση της συνάρτησης. Χρησιμοποιώντας πάλι ενδεικτικά την παραπάνω γραμμική συνάρτηση,

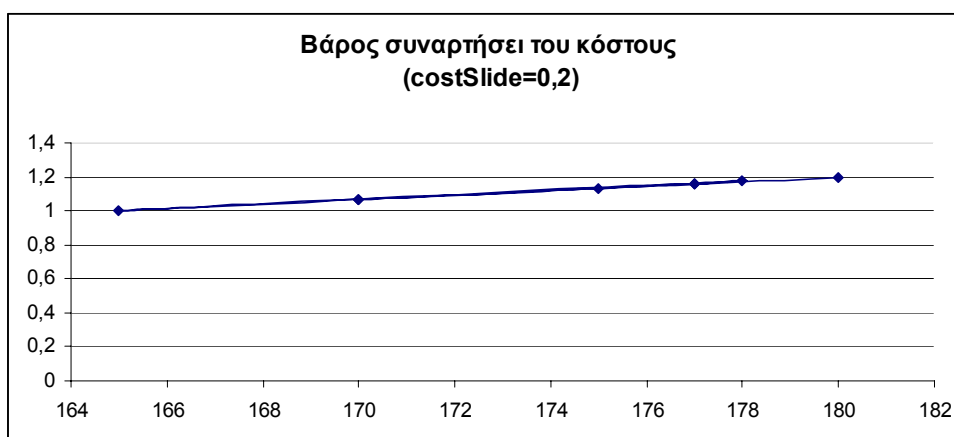
στην περίπτωση του κόστους παρουσιάζονται στη συνέχεια οι μεταβολές για δύο διαφορετικές τιμές του CostSlope στα Σχήματα 40 και 41:



Σχήμα 39: Γραφική αναπαράσταση μεταβολής κόστους



Σχήμα 40: Γραφική αναπαράσταση μεταβολής κόστους για συντελεστή βάρους ίσο με 0.8



Σχήμα 41: Γραφική αναπαράσταση μεταβολής κόστους για συντελεστή βάρους ίσο με 0.2

Η σελίδα αυτή είναι σκόπιμα λευκή

Από τα παραπάνω φαίνεται ότι όταν το *CostSlope* είναι μικρότερο, δηλαδή το ενδιαφέρον για το κόστος είναι μικρό, τότε το βάρος που αντιστοιχεί σε κάθε κόστος είναι μικρότερο από αυτό που αντιστοιχεί, όταν το ενδιαφέρον για το κόστος είναι μεγαλύτερο. Αντίστοιχα ισχύουν και για το χρόνο και τη διαθεσιμότητα.

- Όπως έχει ήδη αναφερθεί οι προτιμήσεις του χρήστη εκφράζονται ως τιμή κλίσης (slope factors) για τις παραμέτρους QoS που αφορούν τη διαθεσιμότητα, το κόστος και το χρόνο τα *CostSlope*, *TimeSlope* και *AvailabilitySlope* θα πρέπει να συνδέονται με κάποια σχέση. Η σχέση αυτή μπορεί να διαφέρει αναλόγως με τη μορφή της συνάρτησης (για παράδειγμα αν είναι της μορφής $y=ax$ ή $y=ax^2$), θα πρέπει όμως σε κάθε περίπτωση να δείχνει ότι αύξηση ενός εκ των μεγεθών προκαλεί μείωση των άλλων. Επομένως:

$$AvailabilitySlope + CostSlope + TimeSlope \leq a \quad (8)$$

όπου a μη μηδενικός θετικός αριθμός. Στο μηχανισμό που αναπτύχθηκε το a τέθηκε ίσο με 3, ενώ θα μπορούσε να είναι κάποιος άλλος μεγαλύτερος αριθμός οπότε οι ακραίες τιμές για τα *CostSlope*, *TimeSlope* και *AvailabilitySlope* θα μεγάλωναν, δίνοντας τη δυνατότητα για ακόμη πιο αυξητικές τάσεις στις συναρτήσεις. Τελικά, πάντως, η επιλογή της τιμής της σταθεράς a μπορεί να διαφέρει ανάλογα με τη μορφή της συνάρτησης f_{Cost} , f_{Time} και $f_{Availability}$ ενώ μπορεί να είναι παραμετρική οριζόμενη από το χρήστη.

- Επιπρόσθετα θα πρέπει η συνάρτηση να λαμβάνει υπόψη και τις διαφορές μεταξύ των τιμών καθώς για παράδειγμα στην περίπτωση του κόστους δεν έχει σημασία κατά την επιλογή των υπηρεσιών το ύψος του κόστους καθαυτό, αλλά η σχέση μεταξύ των κοστών των προσφορών. Αντίστοιχα ισχύουν και για το χρόνο εκτέλεσης και τη διαθεσιμότητα. Αυτό επιτυγχάνεται με την αντικατάσταση της μεταβλητής της συνάρτησης, με τον τύπο:

$$x' = \frac{x - \min}{\max - \min} \quad (9)$$

όπου $\max$ είναι το μέγιστο κόστος, χρόνος ή διαθεσιμότητα που συναντάται στις προσφορές για την υπηρεσία που εξετάζεται και $\min$ το ελάχιστο. Υπάρχει επιπλέον η επιλογή να προστεθεί στον παρονομαστή ένας πολύ μικρός θετικός όρος που θα προστίθεται έτσι ώστε να αποφευχθεί ο μηδενισμός του παρονομαστή στην περίπτωση που $\max$ και $\min$ είναι το ίδιο (κάτι που βέβαια είναι πολύ δύσκολο να συμβεί). Επίσης, επειδή κατά τον υπολογισμό του βάρους που αντιστοιχεί στην ελάχιστη διαθεσιμότητα, μηδενίζεται ο αριθμητής και επομένως όλοι οι όροι της συνάρτησης που περιλαμβάνουν τον ανωτέρω τύπο, θα πρέπει να υπάρχει και ένας σταθερός όρος στη συνάρτηση έτσι ώστε να αποφεύγεται ο μηδενισμός της και στη συνέχεια κατά τον υπολογισμό του τροποποιημένου λόγου κόστους και χρόνου / διαθεσιμότητα να αποφεύγεται η διαίρεση με το 0. Τα παραπάνω αποφεύγονται καθώς η τελικές συναρτήσεις προκύπτουν εκθετικά των παραπάνω μεγεθών όπως παρουσιάζονται αναλυτικά στη συνέχεια:

$$F_{Availability}(x) = \exp\left(AvailabilitySlope * \frac{x - MinAvailabilityValue}{MaxAvailabilityValue - MinAvailabilityValue}\right) \quad (10)$$

$$F_{Cost}(x) = \exp\left(CostSlope * \frac{x - MinCostValue}{MaxCostValue - MinCostValue}\right) \quad (11)$$

$$F_{Time}(x) = \exp\left(TimeSlope * \frac{x - MinTimeValue}{MaxTimeValue - MinTimeValue}\right) \quad (12)$$

όπου x είναι η τιμή της διαθεσιμότητας, του κόστους και του χρόνου για κάθε υπηρεσία και $MinCostValue$, $MinAvailabilityValue$, $MinTimeValue$, $MaxCostValue$,

MaxAvailabilityValue και MaxTimeValue είναι οι ελάχιστες και οι μέγιστες τιμές αυτών των τριών ενδεικτικών παραμέτρων

Οι παραπάνω συναρτήσεις σύμφωνα με όσα αναφέρθηκαν αποδεικνύεται ότι δεν είναι μοναδικές. Η μορφή των συναρτήσεων και οι τιμές των σταθερών όρων που αυτές περιέχουν μπορούν να αποτελέσουν περαιτέρω αντικείμενο έρευνας. Προέκυψαν ως αποτέλεσμα έρευνας και δοκιμών και λειτουργούν παρέχοντας ικανοποιητικά αποτελέσματα όπως παρουσιάζονται στο κεφάλαιο 7.

6.4 Καινοτόμος Αλγόριθμος Απεικόνισης Δυναμικών Ροών Εργασίας και Επιλογής Υπηρεσιών για Διαχείριση της Ποιότητας Υπηρεσίας

Σε αυτή την ενότητα της εργασίας μας, περιγράφουμε τον αλγόριθμο που χρησιμοποιείται στο μηχανισμό απεικόνισης των δυναμικών ροών εργασίας προκειμένου να επιλεγθούν οι υπηρεσίες ανά τύπο υπηρεσίας με βάση τις παραμέτρους ποιότητας υπηρεσίας και τη ροή εργασίας της εφαρμογής (application workflow). Ο κύριος στόχος του αλγορίθμου είναι να καταλήξει σε βέλτιστη επιλογή σε σχέση με τις προτιμήσεις του χρήστη και τη προσφερόμενη ποιότητα από τους παρόχους των υπηρεσιών. Κατ' αρχήν, η ροή εργασίας της εφαρμογής μετατρέπεται σε αφηρημένη ροή εργασίας (abstract workflow) με βάση τη σημασία κάθε τύπου υπηρεσίας που περιλαμβάνεται στη ροή εργασίας. Κατά δεύτερο, οι υποψήφιες υπηρεσίες ανά τύπο υπηρεσίας ταξινομούνται με βάση τις παραμέτρους QoS και τις προτιμήσεις του χρήστη. Έτσι προκύπτει μια ταξινομημένη λίστα με τις "καλύτερες προσφορές" υπηρεσιών ανά τύπο υπηρεσίας και, με βάση αυτή τη λίστα μπορεί να γίνουν αντικαταστάσεις στην αρχική ροή εργασίας προκειμένου να επιτευχθεί η βέλτιστη

απεικόνιση. Στη συνέχεια, εφαρμόζεται ένας μηχανισμός χρονοπρογραμματισμού είτε αν ο χρήστης έχει δηλώσει προτίμηση βελτιστοποίησης χρόνου (όπως δηλώνεται στην ενότητα 6.2 ή προκειμένου να ικανοποιηθεί ο χρονικός περιορισμός (προθεσμία) που έχει οριστεί από το χρήστη.

Ειδικότερα, ο αλγόριθμος αποτελείται από πέντε (5) κύρια βήματα, που περιγράφονται καταρχήν ενώ ακολουθεί λεπτομερής επεξήγηση κάθε βήματος:

- Βήμα 1: Υπολογισμός βοηθητικών τιμών (πilotικές τιμές για την ολοκλήρωση του αλγορίθμου) με βάση τις αρχικά ληφθείσες τιμές των παραμέτρων QoS για τις διαθέσιμες υποστάσεις υπηρεσιών.
- Βήμα 2: Αρχική απεικόνιση ροής εργασίας σε σχέση με τους περιορισμούς και τις προτιμήσεις που έχουν οριστεί από το χρήστη.
- Βήμα 3: Αποδοχή υποψήφιας υπηρεσίας για κάθε τύπο υπηρεσίας με βάση τις αρχικές παραμέτρους και υπολογισμός πρόσθετου κόστους προκειμένου να αυξηθεί η διαθεσιμότητα. Το δεύτερο μπορεί να θεωρηθεί ως βήμα βελτιστοποίησης προκειμένου να επιτευχθούν καλύτερα επίπεδα διαθεσιμότητας ενώ δεν παραβιάζονται οι περιορισμοί κόστους που έχει ορίσει ο χρήστης.
- Βήμα 4: Δημιουργία λίστας με τις "καλύτερες προσφορές" για την τελική συγκεκριμένη ροή εργασίας προκειμένου να βρεθούν τυχόν αντικαταστάσεις με στόχο να επιτευχθεί η βέλτιστη επιλογή και εκ νέου υπολογισμός της συνολικής διαθεσιμότητας, κόστους και χρόνου εκτέλεσης της ροής εργασίας.
- Βήμα 5: Εφαρμογή μηχανισμού χρονοπρογραμματισμού που επιτρέπει την επιλογή περισσότερων από μιας υποστάσεων υπηρεσιών ανά τύπο υπηρεσίας προκειμένου να ικανοποιηθούν οι χρονικοί περιορισμοί που έχουν τεθεί από το χρήστη. Στον ενσωματωμένο μηχανισμό χρονοπρογραμματισμού γίνεται η παραδοχή ότι, κάθε φορά που διαιρείται μια υπηρεσία, ο αντίστοιχος χρόνος εκτέλεσης και το κόστος μειώνονται

στο μισό. Ωστόσο οι προσομοιώσεις έδειξαν ότι ο χρόνος και το κόστος δεν μειώνονται στο μισό στο σύνολό του αλλά προστίθεται μια ποινή (σαν μια σταθερή χρέωση - πάγιο). Ο μηχανισμός χρονοπρογραμματισμού έχει σχεδιαστεί ώστε να λαμβάνει υπόψη και αυτή την χρέωση.

Στο αλγόριθμο απεικόνισης χρησιμοποιούνται οι συναρτήσεις που αναλύθηκαν στην προηγούμενη παράγραφο της παρούσης διατριβής. Για τις προτιμήσεις του χρήστη χρησιμοποιούνται οι συντελεστές κλίσης που μπορούν να πάρουν τιμές στην περιοχή $0 \leq \text{κλίση} \leq 1$. Όσο υψηλότερη είναι η τιμή της κλίσης τόσο σημαντικότερη είναι η αντίστοιχη παράμετρος για το χρήστη. Τα δεκαδικά ψηφία στην τιμή της κλίσης επιτρέπουν τη λεπτομερή εισαγωγή των προτιμήσεων του χρήστη. Επομένως η εξίσωση (8) γίνεται:

$$\text{AvailabilitySlope} + \text{CostSlope} + \text{TimeSlope} \leq 3 \quad (13)$$

Στη συνέχεια, περιγράφονται λεπτομερώς τα κύρια βήματα του αλγορίθμου μαζί με τα υποβήματά τους:

Βήμα 1: Υπολογισμός βοηθητικών τιμών (πilotικών τιμών για την ολοκλήρωση του αλγορίθμου)

- i. Υπολογισμός ελάχιστων και μέγιστων τιμών διαθεσιμότητας, κόστους και χρόνου για κάθε συγκεκριμένο τύπο υπηρεσίας, λαμβάνοντας υπόψη όλες τις διαθέσιμες υποψήφιες υπηρεσίες.
- ii. Υπολογισμός των pilotικών τιμών για τις παραμέτρους διαθεσιμότητας, κόστους και χρόνου με βάση τις ελάχιστες και τις μέγιστες τιμές τους με χρήση των συναρτήσεων (10), (11) και (12).
- iii. Υπολογισμός των νέων τιμών των παραμέτρων που θα χρησιμοποιηθούν στη συνέχεια με βάση τις παραπάνω συναρτήσεις ως εξής:

$$NewCostValue = InitialCostValue * F_{Cost}(InitialCostValue) \quad (14)$$

$$NewAvailabilityValue = InitialAvailabilityValue * F_{Availability}(InitialAvailabilityValue) \quad (15)$$

$$NewTimeValue = InitialTimeValue * F_{Time}(InitialTimeValue) \quad (16)$$

Στις παραπάνω εξισώσεις, οι μεταβλητές *InitialCostValue*, *InitialTimeValue* και *InitialAvailabilityValue* είναι οι τιμές των παραμέτρων κόστους, χρόνου και διαθεσιμότητας αντίστοιχα που ελήφθησαν αρχικά από τους παρόχους για κάθε υπηρεσία.

- iv. Υπολογισμός του παρακάτω δείκτη μετατροπής που θα χρησιμοποιηθεί στη συνέχεια προκειμένου να γίνουν οι επιλογές:

$$ConvertedIndex = \frac{NewCostValue * NewTimeValue}{NewAvailabilityValue} \quad (17)$$

Βήμα 2: Αρχική απεικόνιση ροής εργασίας

- i. Με βάση τις παραμέτρους που ορίζονται από το χρήστη (περιορισμοί) και τις προτιμήσεις, για κάθε τύπο υπηρεσίας, επιλέγεται η υπηρεσία που ικανοποιεί το ελάχιστο απαιτούμενο επίπεδο διαθεσιμότητας (αν ο χρήστης έχει ορίσει όριο διαθεσιμότητας) με το ελάχιστο κόστος.
- ii. Αν η τιμή συνολικού κόστους, που υπολογίζεται για όλες τις επιλεγμένες υπηρεσίες υπερβαίνει το όριο που έχει θέσει ο χρήστης για το συνολικό κόστος, η διαδικασία απεικόνισης της ροής εργασίας τερματίζεται και δεν μπορεί να ολοκληρωθεί η επιλογή με βάση τους περιορισμούς του χρήστη. Διαφορετικά, ο αλγόριθμος συνεχίζει στο επόμενο βήμα.

Βήμα 3: Αποδοχή υποψηφίας υπηρεσίας για κάθε τύπο υπηρεσίας με βάση τις αρχικές παραμέτρους QoS

- i. Για κάθε τύπο υπηρεσίας, καθορίζεται η "βέλτιστη" υπηρεσία από αυτή που επιλέχθηκε στο Βήμα 2. Το "βέλτιστο" ορίζεται από τη μικρότερη τιμή του δείκτη *ConvertedIndex* (όπως προκύπτει από την εξίσωση 17).
- ii. Αν για κάποιο τύπο υπηρεσίας δεν υπάρχει "βέλτιστη" δυνατή αντικατάσταση, αυτός ο τύπος υπηρεσίας εξαιρείται από την υπόλοιπη εκτέλεση του αλγορίθμου εφόσον δεν μπορεί να γίνει καμία βελτιστοποίηση ενώ, αν συμβεί αυτό για όλους τους τύπους υπηρεσιών της ροής εργασίας, ο αλγόριθμος τερματίζεται και η αρχική απεικόνιση της ροής εργασίας θεωρείται και τελική. Διαφορετικά, ο αλγόριθμος συνεχίζει στο επόμενο υπο-βήμα.
- iii. Επιλογή των υποψηφίων υπηρεσιών για κάθε τύπο υπηρεσίας με τη μικρότερη τιμή δείκτη *ConvertedIndex* (όπως προκύπτει από την εξίσωση 17).
- iv. Υπολογισμός των διαφορών στις τιμές διαθεσιμότητας, κόστους και χρόνου μεταξύ της αρχικής επιλογής υπόστασης υπηρεσίας και της νέας που την αντικατέστησε.

Βήμα 4: Δημιουργία καταλόγου με τις "καλύτερες προσφορές" για την τελική συγκεκριμένη ροή εργασίας προκειμένου να βρεθούν δυνατές αντικαταστάσεις

- i. Για κάθε διαφορά που υπολογίζεται στο Βήμα 3, υπολογίζεται εκ νέου ο δείκτης μετατροπής *ConvertedIndex* (όπως προκύπτει από την εξίσωση 17). Βασικά, εκτελείται ξανά το Βήμα 1 του αλγορίθμου θεωρώντας ως αρχικές τιμές για τις υπηρεσίες τις παραπάνω διαφορές και οι αντικαταστάσεις γίνονται με βάση τις διαφορές τους (για παράδειγμα είναι προτιμότερο να δαπανηθούν 6 μονάδες κόστους προκειμένου να αυξηθεί η διαθεσιμότητα κατά 4% από το να δαπανηθούν 10 μονάδες κόστους προκειμένου να αυξηθεί κατά 5%).

- ii. Επιλέγονται οι υποστάσεις υπηρεσιών με το μικρότερο δείκτη μετατροπής *ConvertedIndex* (όπως προκύπτει από την εξίσωση 17).
- iii. Με βάση τη νέα επιλογή υποστάσεων υπηρεσιών, υπολογίζονται ξανά οι τιμές συνολικού κόστους και χρόνου εκτέλεσης της ροής εργασίας.
- iv. Αν ικανοποιείται ο περιορισμός που έχει θέσει ο χρήστης για το κόστος, πραγματοποιείται η αντικατάσταση της νέας επιλογής. Διαφορετικά, η νέα επιλογή εξαιρείται από τις διαθέσιμες υποστάσεις υπηρεσιών. Όσο αφορά το χρονικό περιορισμό (πιθανή προθεσμία που έχει οριστεί από το χρήστη), αναλύεται στο Βήμα 5 – μηχανισμός χρονοπρογραμματισμού – που περιγράφεται στη συνέχεια.
- v. Ο αλγόριθμος επαναλαμβάνεται και συνεχίζει από το Βήμα 3 για όλους τους τύπους υπηρεσιών. Αφού εκτελεστεί ο αλγόριθμος για όλους τους τύπους υπηρεσιών, προχωρά στο επόμενο βήμα.

Βήμα 5: Μηχανισμός χρονοπρογραμματισμού

- i. Κατηγοριοποίηση των τύπων υπηρεσιών (με αποτέλεσμα τιμές προτεραιότητας) με βάση τις υποδείξεις της εφαρμογής που έχουν οριστεί μέσα στη ροή εργασίας της εφαρμογής (application workflow). Για κάθε διαφορετική προτεραιότητα, δημιουργείται μια νέα κατηγορία όπου ανήκουν όλοι οι τύποι υπηρεσιών με την καθορισμένη προτεραιότητα και, λόγω αυτού, αυτοί οι τύποι υπηρεσιών μπορούν να εκτελούνται παράλληλα. Στη συνέχεια, υπολογίζεται ο χρόνος που απαιτείται για να εκτελεστεί ολόκληρη η ροή εργασίας. Από τα παραπάνω γίνεται σαφές ότι δεν είναι σημαντικός ο χρόνος ολοκλήρωσης ενός τύπου υπηρεσίας αλλά ο χρόνος ολοκλήρωσης της κατηγορίας όπου ανήκει ο συγκεκριμένος τύπος υπηρεσίας.
- ii. Αν ο χρήστης έχει ορίσει υψηλότερη προτίμηση για τη διαθεσιμότητα ή για το κόστος από την αντίστοιχη για το χρόνο εκτέλεσης και εφόσον πληρείται ο χρονικός

- περιορισμός, το βήμα χρονοπρογραμματισμού ολοκληρώνεται. Διαφορετικά, ο αλγόριθμος συνεχίζει στα επόμενα υπο-βήματα.
- iii. Καθορισμός του τύπου υπηρεσίας που επηρεάζει σημαντικά το χρόνο (time-critical), ως αυτός για τον οποίο ο χρόνος εκτέλεσης της προεπιλεγμένης υποψήφιας υπηρεσίας είναι μεγαλύτερος από το χρόνο εκτέλεσης άλλης υποψήφιας υπηρεσίας διαφορετικού τύπου υπηρεσίας. Προκειμένου να γίνει χρονοπρογραμματισμός αυτού του τύπου υπηρεσίας που επηρεάζει σημαντικά το χρόνο θα πρέπει να υπάρχει διαθέσιμη τουλάχιστον μια υποψήφια υπηρεσία, καθώς ορισμένες υποψήφιες μπορεί να έχουν εξαιρεθεί κατά την εκτέλεση των πρώτων τεσσάρων βημάτων του αλγορίθμου όπως περιγράφεται παραπάνω. Αν υπάρχει άλλος τύπος υπηρεσίας που επηρεάζει σημαντικά το χρόνο, η κατηγορία στην οποία ανήκει και η πιο χρονοβόρα υποψήφια υπηρεσία αποθηκεύονται. Αν δεν βρεθεί τύπος υπηρεσίας που να επηρεάζει σημαντικά το χρόνο, αυτό το υπο-βήμα ολοκληρώνεται. Διαφορετικά, ο μηχανισμός χρονοπρογραμματισμού συνεχίζεται.
- iv. Καθορισμός της "βέλτιστη" υποψήφιας υπηρεσίας για τον παραπάνω τύπο υπηρεσίας. Ως "βέλτιστη" θεωρείται η υπόσταση υπηρεσίας που έχει το μικρότερο χρόνο εκτέλεσης και ονομάζεται *Πρωτεύουσα Υποψήφια (Primary Candidate)*. Αυτή η υποψήφια δεν πρέπει να έχει επιλεγεί προηγουμένως και δεν πρέπει να έχει απενεργοποιηθεί – όπως μπορεί να συμβεί στο επόμενο υπο-βήμα.
- v. Ελέγχος αν η προσθήκη της νέας υπόστασης υπηρεσίας αυξάνει ή μειώνει το χρόνο εκτέλεσης του τύπου υπηρεσίας που επηρεάζει σημαντικά το χρόνο. Αν αυξάνει το χρόνο, απενεργοποιείται και ο μηχανισμός χρονοπρογραμματισμού επιστρέφει στο υπο-βήμα (i). Διαφορετικά, ο αλγόριθμος συνεχίζει στα επόμενα βήματα.
- vi. Ελέγχος αν η προσθήκη της νέας υπηρεσίας ελαττώνει το χρόνο εκτέλεσης του τύπου υπηρεσίας ώστε να είναι ακόμα μικρότερος από το χρόνο εκτέλεσης άλλου τύπου

- υπηρεσίας της ίδιας κατηγορίας. Για κάθε τύπο υπηρεσίας που συμβαίνει αυτό, ο μηχανισμός χρονοπρογραμματισμού προχωρά στο επόμενο υπο-βήμα (υπο-βήμα (vii)). Διαφορετικά, προχωρά στο υπο-βήμα (viii).
- vii. Για το δευτερεύοντα τύπο υπηρεσίας που εξετάζεται, ορίζεται ξανά η "βέλτιστη" υπηρεσία εφαρμόζοντας το υπο-βήμα (iii). Αν δεν βρεθεί αυτή η υπηρεσία, ο αλγόριθμος συνεχίζει με τον επόμενο δευτερεύοντα τύπο υπηρεσίας. Διαφορετικά, αν βρεθεί υπηρεσία εξετάζεται αν η προσθήκη αυτή αυξάνει ή ελαττώνει το χρόνο εκτέλεσης του δευτερεύοντα τύπου υπηρεσίας στον οποίο ανήκει. Αν αυξάνει το χρόνο, η υπηρεσία απενεργοποιείται και ο αλγόριθμος συνεχίζει με τον επόμενο δευτερεύοντα τύπο υπηρεσίας. Αν μειώνει το χρόνο, ανήκει στις καλούμενες *Δευτερεύουσες Υποψήφιες Υπηρεσίες* (*Secondary Service Instances*). Αφού ολοκληρωθεί αυτό το βήμα για καθένα από τους τύπους υπηρεσιών που ορίστηκαν στο υπο-βήμα (v), ο μηχανισμός χρονοπρογραμματισμού προχωρά στο υπο-βήμα (viii).
- viii. Υπολογισμός του συνολικού κόστους της ροής εργασίας μετά την προσθήκη των δευτερευουσών υποψηφίων υπηρεσιών. Αν δεν παραβιάζεται ο περιορισμός κόστους που έχει θέσει ο χρήστης, επιβεβαιώνονται οι αλλαγές και επιλέγονται προσθετικά και οι δευτερεύουσες υποψήφιες. Διαφορετικά, η πρωτεύουσα υποψήφια απενεργοποιείται και ο μηχανισμός χρονοπρογραμματισμού επιστρέφει στο υπο-βήμα (i).

6.5 Καινοτόμο Σχήμα Καθορισμού Περισσοτέρων

Συγκεκριμένων Ροών Εργασίας

Στην παρούσα ενότητα παρουσιάζεται ένα σχήμα που επιτρέπει τη δημιουργία εναλλακτικών συγκεκριμένων ροών εργασίας (για μια δεδομένη ροή εργασίας εφαρμογής) που συναντούν το επιθυμητό επίπεδο ποιότητας υπηρεσίας, όπως αυτό εκφράζεται με τους καθορισμένους από το χρήστη περιορισμούς και τις προτιμήσεις αυτού (επίσης παρουσιάζεται στην αναφορά [63]). Οι ροές εργασίας είναι διαφορετικές από την άποψη της παρεχόμενης ποιότητας και βασίζονται στις διάφορες επιλογές που ο τελικός χρήστης λαμβάνει από ένα σύνολο παρόχων για την εκτέλεση της ροής εργασίας.

Σημαντικός παράγοντας για τον καθορισμό των διαφορετικών συγκεκριμένων ροών εργασίας είναι το χρονικό πλαίσιο στο οποίο θα εκτελεστούν. Αναγνωρίζοντας ότι οι πάροχοι δημοσιεύουν διαφορετικές προσφορές βασισμένες στην τρέχουσα διαθεσιμότητα και το φόρτο εργασίας τους, αυτές οι προσφορές διαφέρουν στο κόστος για την εκτέλεση μια συγκεκριμένης ροής εργασίας. Η διαφορά κόστους που προκύπτει από την εκτέλεση των ροών σε διαφορετικά χρονικά πλαίσια μπορεί είτε να επιστραφεί στο χρήστη είτε να χρησιμοποιηθεί προκειμένου να παρασχεθεί ένα πιο υψηλό επίπεδο ποιότητας υπηρεσιών. Στη συνέχεια παρουσιάζεται η αρχιτεκτονική του σχήματος, τα δομικά στοιχεία αυτής καθώς και η λειτουργικότητα που προσφέρουν. Επιπλέον θα πρέπει να σημειωθεί ότι τα αποτελέσματα προσομοίωσης του σχήματος παρουσιάζονται στην ενότητα 7.3.3 της παρούσης διατριβής.

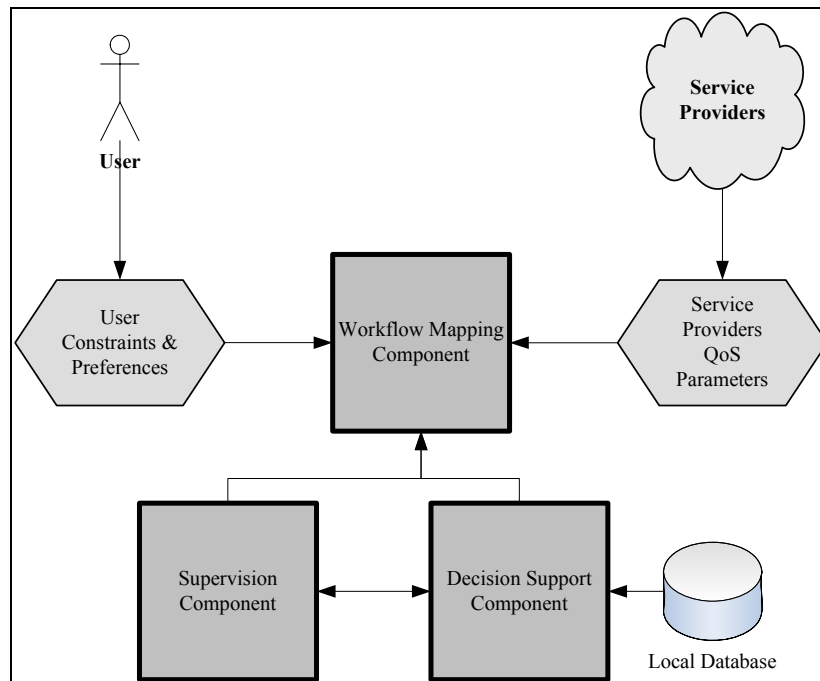
6.5.1 Αρχιτεκτονική

Το σχήμα που επιτρέπει τον καθορισμό εναλλακτικών δυναμικών ροών εργασίας αρχικοποιείται από τον χρήστη, που υποβάλει προς εκτέλεση μια ροή εργασίας σε περιβάλλον πλέγματος. Η εκτέλεση της ροής εργασίας με επιλογή των αντίστοιχων υπηρεσιών πλέγματος θα πρέπει να γίνει με γνώμονα το επιθυμητό επίπεδο της ποιότητας υπηρεσιών. Με βάση τα ανωτέρω, το συγκεκριμένο σχήμα είναι σε θέση να καθορίζει διαφορετικές συγκεκριμένες ροές εργασίας (concrete workflows) και να επιτρέπει στο χρήστη την επιλογή μιας εξ' αυτών και να συνεχίσει με την εκτέλεσή της. Τα δομικά στοιχεία της αρχιτεκτονικής που καθιστά εφικτή την παραπάνω δυνατότητα εμφανίζονται στο σχήμα της επόμενης σελίδας (Σχήμα 42) και αναλύονται ακολούθως.

6.5.1.1 Στοιχείο Διαχείρισης

Το στοιχείο Διαχείρισης (Supervision Component) χρησιμοποιείται για τη διαχείριση, οδηγία, επίβλεψη, έλεγχο και έναρξη της συγκεκριμένης διαδικασίας καθορισμού ροών εργασίας. Για μια δεδομένη ροή εργασίας εφαρμογής και μια χρονική παράμετρο εκτέλεσης (ως περιορισμός που τίθεται από το χρήστη) απαιτείται ένα στοιχείο προκειμένου να καθοριστούν τα διάφορα χρονικά πλαίσια εκτέλεσης της ροής εργασίας. Μόλις ο χρήστης θέτει τους περιορισμούς ποιότητας και τις προτιμήσεις του, το στοιχείο διαχείρισης αρχικοποιεί τη διαδικασία καθορισμού συγκεκριμένων ροών εργασίας (concrete workflows) που παρέχουν διαφορετικά επίπεδα ποιότητας.

Ο κύριος ρόλος του στοιχείου αυτού είναι η ενορχήστρωση της λειτουργία του μηχανισμού απεικόνισης δυναμικών ροών εργασίας για την εκτέλεσή τους στα διαφορετικά χρονικά πλαίσια.



Σχήμα 42: Αρχιτεκτονική συστήματος

6.5.1.2 Στοιχείο Υποστήριξης Αποφάσεων

Το στοιχείο Υποστήριξης Αποφάσεων (Decision Support Component) χρησιμοποιείται για να παρέχει τις ιστορικές πληροφορίες σχετικά με τη συμπεριφορά των φορέων παροχής υπηρεσιών για τις προηγούμενες υποβληθείσες εργασίες. Οι προαναφερθείσες ιστορικές πληροφορίες περιγράφονται με την παράμετρο «*Characterization*» η οποία αντικαθιστά την παράμετρο «*Availability*» (διαθεσιμότητα) στο μηχανισμό απεικόνισης δυναμικών ροών εργασίας (βλ. ενότητα 6.2) και με τον τρόπο αυτό προσθέτει ένα επιπλέον κριτήριο κατά τη διαδικασία απεικόνισης των ροών και επιλογής των υπηρεσιών πλέγματος. Περαιτέρω περιγράφουμε τον τρόπο υπολογισμού της παραπάνω παραμέτρου:

- i. Μια τοπική βάση δεδομένων χρησιμοποιείται προκειμένου να αποθηκευτούν οι ιστορικές πληροφορίες για τους φορείς παροχής υπηρεσιών. Αυτές οι πληροφορίες αναφέρονται σε προηγούμενες υπογεγραμμένες συμβάσεις παροχής υπηρεσιών και

Η σελίδα αυτή είναι σκόπιμα λευκή

στον αριθμό των παραβιάσεων αυτών για κάθε φορέα παροχής υπηρεσιών. Επιπλέον, η τοπική βάση δεδομένων του τμήματος υποστήριξης αποφάσεων αποθηκεύει τις πληροφορίες σχετικά με την προηγούμενη εκτέλεση εργασιών ανά φορέα παροχής υπηρεσιών μαζί με τις πληροφορίες ποιότητάς τους (τιμές χρόνου εκτέλεσης και κόστους).

- ii. Η τιμή της αξιοπιστίας (Reliability) για κάθε φορέα παροχής υπηρεσιών υπολογίζεται βάσει του ποσοστού των παραβιασμένων συμβολαίων SLAs σε σύγκριση με τον αριθμό των υπογεγραμμένων:

$$Reliability = 1 - \frac{NumberOfViolatedSLAs}{NumberOfSignedSLAs} \quad (18)$$

- iii. Η παράμετρος που χαρακτηρίζει κάθε φορέα παροχής υπηρεσιών («*Characterization*») παράγεται ως πολλαπλασιασμός της αξιοπιστίας και των τιμών διαθεσιμότητας:

$$Characterisation = Reliability * Availability \quad (19)$$

Ο πολλαπλασιασμός της αξιοπιστίας με τη διαθεσιμότητα για κάθε φορέα παροχής υπηρεσιών γίνεται ώστε ακόμα κι αν ένας φορέας παροχής υπηρεσιών είναι αξιόπιστος αλλά μη διαθέσιμος να μην μπορεί να επιλεγεί για μια εκτέλεση εργασίας και αντίστροφα στην περίπτωση όπου ένας φορέας παροχής υπηρεσιών είναι διαθέσιμος αλλά μη αξιόπιστος. Επιπλέον, εάν οι ιστορικές πληροφορίες για έναν φορέα παροχής υπηρεσιών δεν είναι διαθέσιμες, η τιμή της παραμέτρου αξιοπιστίας τίθεται ίση με την μονάδα ("1") και η διαδικασία επιλογής σε αυτήν την περίπτωση γίνεται αποκλειστικά βάσει της διαθεσιμότητά του παρόχου.

6.5.1.3 Στοιχείο Απεικόνισης Δυναμικών Ροών Εργασίας

Πρόκειται για το μηχανισμό απεικόνισης δυναμικών ροών εργασίας που περιγράφεται στην ενότητα 6.2 της διατριβής.

6.5.2 Λειτουργικότητα

Στη συνέχεια περιγράφεται η λειτουργικότητα που παρέχεται από το συγκεκριμένο σχήμα και η προστιθέμενη αξία που προκύπτει από τη συνδυασμένη χρήση των επιμέρους στοιχείων σε αντίθεση με τη μεμονωμένη χρήση αυτών. Στη συνέχεια περιγράφεται η ροή της λογικής μεταξύ των στοιχείων του σχήματος (και παρατίθεται το αντίστοιχο διάγραμμα):

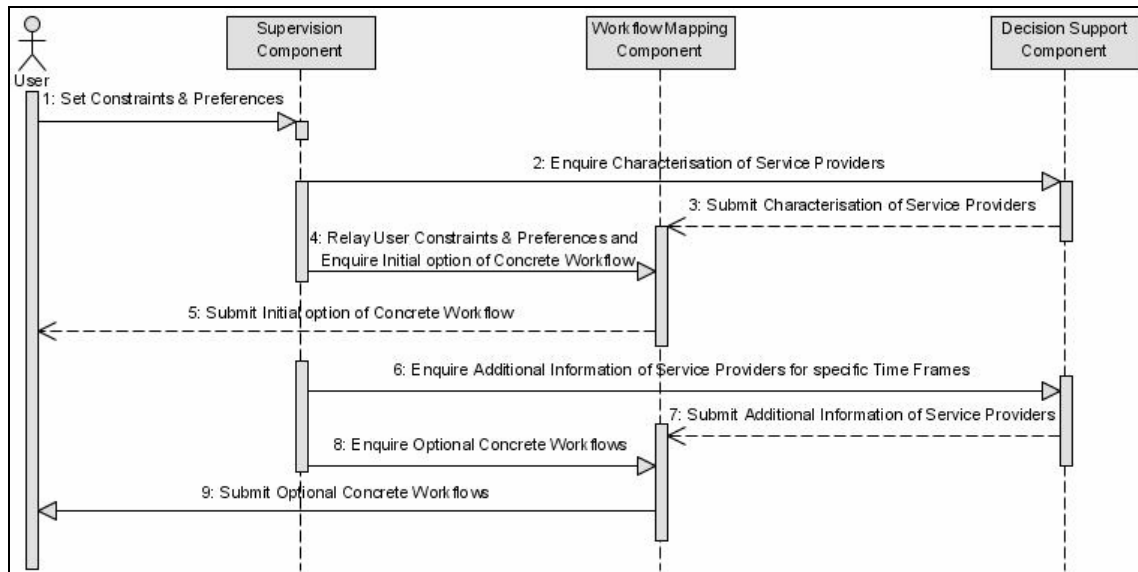
1. Καθορισμός περιορισμών & προτιμήσεων: Ο χρήστης που αρχικοποιεί μια ροή εργασίας εκφράζει τις απαιτήσεις του αναφορικά με το επίπεδο της ποιότητας υπηρεσιών. Αυτές οι απαιτήσεις εκφράζονται υπό μορφή παραμέτρων ποιότητας, για τις οποίες ο χρήστης θέτει τους περιορισμούς (για παράδειγμα κατώτατα όρια/περιορισμοί στις συγκεκριμένες τιμές των παραμέτρων ποιότητας) μαζί με τις πιθανές προτιμήσεις, οι οποίες συνδέουν μια ιδιότητα βάρους με τις παραμέτρους (βλ. ενότητα 0).
2. Ερώτηση για την τιμή χαρακτηρισμού (Characterization) των φορέων παροχής υπηρεσιών: Δεδομένου ότι οι περιορισμοί και οι προτιμήσεις των χρηστών αντιμετωπίζονται αρχικά από το στοιχείο διαχείρισης, το τελευταίο ερευνά τις τιμές χαρακτηρισμού των φορέων παροχής υπηρεσιών από το τμήμα υποστήριξης αποφάσεων βασισμένο στη λειτουργία που ζητείται από το χρήστη μέσω της ροής εργασίας της εφαρμογής (application workflow).
3. Απάντηση για το χαρακτηρισμό των φορέων παροχής υπηρεσιών: Το στοιχείο υποστήριξης αποφάσεων απαντά με την αποστολή των τιμών χαρακτηρισμού στο στοιχείο απεικόνισης δυναμικών ροών εργασίας και επιλογής υπηρεσιών.

4. Αίτηση καθορισμού της αρχικής έκδοσης συγκεκριμένης ροής εργασίας: Το στοιχείο διαχείρισης διοχετεύει τους περιορισμούς και τις προτιμήσεις του χρήστη στο στοιχείο απεικόνισης της ροής εργασίας και επιλογής υπηρεσιών και ζητά τον καθορισμό της αρχικής έκδοσης συγκεκριμένης ροής εργασίας - που αναφέρεται σε μια ροή εργασίας της οποίας η εκτέλεση θα ξεκινήσει άμεσα.
5. Υποβολή της αρχικής έκδοσης συγκεκριμένης ροής εργασίας: Το στοιχείο απεικόνισης ροών εργασίας χρησιμοποιεί τους περιορισμούς και τις προτιμήσεις που τίθενται από το χρήστη μαζί με τις πληροφορίες ποιότητας για τις υπηρεσίες πλέγματος και καθορίζει την αρχική έκδοση της συγκεκριμένης ροής εργασίας.
6. Αίτηση πρόσθετων πληροφοριών για τους φορείς παροχής υπηρεσιών για τα διαφορετικά χρονικά πλαίσια: Το στοιχείο διαχείρισης καθορίζει τα διάφορα χρονικά πλαίσια στα οποία η ροή εργασίας μπορεί να εκτελεσθεί. Για αυτά τα χρονικά πλαίσια, οι πρόσθετες πληροφορίες απαιτούνται (προηγούμενες εκτελέσεις εργασίας σε έναν φορέα παροχής υπηρεσιών σχετικά με το χρόνο εκτέλεσης και το κόστος), οι οποίες λαμβάνονται από το τμήμα υποστήριξης αποφάσεων. Πρέπει να σημειωθεί ότι αυτές οι πληροφορίες μπορούν να ληφθούν από συμβόλαια SLA εάν οι φορείς παροχής υπηρεσιών τοποθετούν επίσης τις πληροφορίες για την εκτέλεση εργασίας στα διαφορετικά χρονικά πλαίσια. Σε αυτήν την περίπτωση η ροή της λογικής παραμένει αμετάβλητη αλλά η πηγή μέσα στο στοιχείο υποστήριξης αποφάσεων δεν είναι η τοπική βάση δεδομένων με τις ιστορικές πληροφορίες αλλά τις τρέχουσες προσφορές που υποβάλλονται από τους φορείς παροχής υπηρεσιών μέσω συμβολαίων SLA.
7. Υποβολή των πρόσθετων πληροφοριών για τους φορείς παροχής υπηρεσιών: Το στοιχείο υποστήριξης αποφάσεων απαντά παρέχοντας τις πρόσθετες πληροφορίες για τους φορείς παροχής υπηρεσιών, οι οποίες αναφέρονται στις προσφορές για

διαφορετικά χρονικά πλαίσια. Αυτές οι πληροφορίες αναλύονται στο στοιχείο απεικόνισης ροών εργασίας.

8. Αίτηση για διαφορετικές εκδόσεις συγκεκριμένων ροών εργασίας: Το στοιχείο διαχείρισης ζητά τον καθορισμό των διαφορετικών εκδόσεων συγκεκριμένων ροών εργασίας, οι οποίες εκτελεστούν σε διαφορετικά χρονικά πλαίσια και όχι άμεσα.
9. Υποβολή των διαφορετικών εκδόσεων συγκεκριμένων ροών εργασίας: Όπως περιγράφεται ήδη, η διαφορά κόστους που προκύπτει από την εκτέλεση των εργασιών σε διαφορετικά χρονικά πλαίσια μπορεί είτε να επιστραφεί στο χρήστη είτε να χρησιμοποιηθεί προκειμένου να παρασχεθεί ένα πιο υψηλό επίπεδο ποιότητας υπηρεσιών, όπως αυτό εκφράζεται με τον καθορισμό των διαφορετικών συγκεκριμένων ροών εργασίας. Βάσει αυτού, το στοιχείο απεικόνισης ροών εργασίας καθορίζει δύο (2) επιπλέον διαφορετικές εκδόσεις συγκεκριμένων ροών. Η πρώτη ικανοποιεί τους περιορισμούς και τις προτιμήσεις που τίθενται από το χρήστη αλλά με το ελαχιστοποιημένο κόστος, δεδομένου ότι η συγκεκριμένη ροή εργασίας θα εκτελεσθεί σε διαφορετικό χρονικό πλαίσιο. Η δεύτερη χρησιμοποιεί το διαθέσιμο κόστος προκειμένου να βελτιστοποιηθεί το προσφερθέν επίπεδο ποιότητας με την παροχή υψηλότερης τιμής για την παράμετρο χαρακτηρισμού χαμηλότερου συνολικού χρόνου εκτέλεσης για το σύνολο της ροής εργασίας χωρίς συγχρόνως να παραβιάσει τους περιορισμούς που τίθενται από το χρήστη.

Η προναφερθείσα λειτουργικότητα παρατίθεται διαγραμματικά στην επόμενη σελίδα (Σχήμα 43):



Σχήμα 43: Ροή λογικής μεταξύ των στοιχείων

6.6 Ενσωμάτωση του Μηχανισμού Απεικόνισης στο μοντέλο

GridVIM

Στο συγκεκριμένο κεφάλαιο (Κεφάλαιο 6) έγινε αναλυτική περιγραφή του μηχανισμού με τον οποίο γίνεται η επιλογή των υπηρεσιών εκείνων, που θα υλοποιήσουν τις αφηρημένες υπηρεσίες μιας αφηρημένης ροής εργασίας. Πλέον τίθεται το θέμα της ενσωμάτωσης του μηχανισμού αυτού στο μοντέλο GridVIM (βλ. ενότητα 4.4.1).

Όπως έχει ήδη αναφερθεί στην ενότητα 4.4.2.4, στο σχεδιασμό του GridVIM περιλαμβάνεται η διεπαφή επιλογής (Selection interface), η οποία αναλαμβάνει την αντιστοίχιση αφηρημένων διεργασιών σε συγκεκριμένες υπηρεσίες με βάση ορισμένους περιορισμούς και απαιτήσεις. Έχει τη δυνατότητα ακόμη και να πραγματοποιεί τη διαδικασία αυτή συνολικά για μια αφηρημένη ροή εργασίας. Για την τελευταία αυτή λειτουργία διαθέτει μια μέθοδο, που δέχεται ως είσοδο μια λίστα δομών ελέγχου, που περιέχει τις αφηρημένες διεργασίες της ροής εργασίας.

Η σελίδα αυτή είναι σκόπιμα λευκή

Η λειτουργία του μηχανισμού απαιτεί εκτός από την αφηρημένη ροή εργασίας και ορισμένες παραμέτρους. Αυτές είναι το κόστος, η διαθεσιμότητα και ο χρόνος εκτέλεσης κάθε υποψήφιας διεργασίας, ένα ανώτατο όριο συνολικού κόστους, ένα κατώτερο όριο διαθεσιμότητας, ένα ανώτατο όριο χρόνου εκτέλεσης για κάθε διεργασία και τις προτιμήσεις του χρήστη για το κόστος, το χρόνο εκτέλεσης και τη διαθεσιμότητα. Επομένως χρειάζονται ορισμένες προσθήκες στο μοντέλο GridVIM:

- Θα πρέπει το προφίλ κάθε υπηρεσίας να περιλαμβάνει και μια παράμετρο υπηρεσίας (*serviceParameter*), που αντιστοιχεί στο κόστος χρήσης της. Επίσης θα πρέπει το προφίλ κάθε υπηρεσίας να περιλαμβάνει μια επιπλέον παράμετρο υπηρεσίας (*serviceParameter*), που θα αντιστοιχεί στη διαθεσιμότητά της, έτσι όπως αυτή υπολογίζεται από τις υπόλοιπες υπηρεσίες της GridVIM και αντίστοιχα μια παράμετρο για το χρόνο εκτέλεσης. Οι παράμετροι αυτές ανήκουν στα μη λειτουργικά χαρακτηριστικά του προφίλ της υπηρεσίας και αποτελούν παραμέτρους της ποιότητας υπηρεσίας.
- Πρέπει επιπλέον να προστεθούν στο προφίλ του συνολικού workflow, οι παράμετροι ανώτατο όριο συνολικού κόστους (*maxCost*), ανώτατο χρονικό όριο εκτέλεσης (*maxTme*), κατώτερο όριο διαθεσιμότητας (*minavailability*), προτίμηση για το κόστος (*costSlope*), προτίμηση για το χρόνο (*timeSlope*) και προτίμηση για τη διαθεσιμότητα (*availabilitySlope*). Όλες αυτές οι παράμετροι ανήκουν επίσης στην ποιότητα υπηρεσίας.

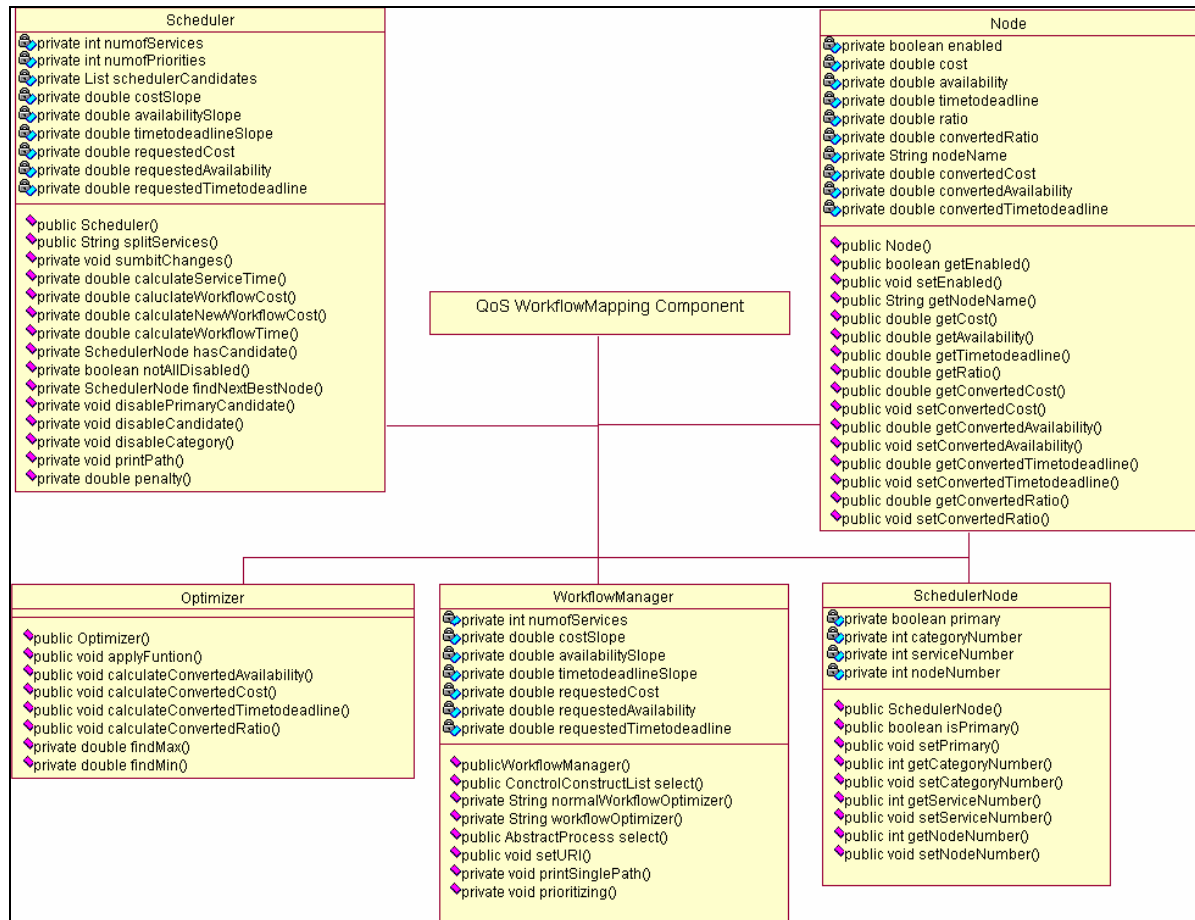
Για την καλύτερη λειτουργία του αλγορίθμου προστέθηκαν στη διεπαφή, ορισμένες νέες μέθοδοι και δημιουργήθηκαν ορισμένες βοηθητικές κλάσεις. Αναλυτικότερα, η κλάση *WorkflowManager* υλοποιεί τη διεπαφή *Selection*. Διαθέτει τις τρεις βασικές μεθόδους της διεπαφής, εκ των οποίων υλοποιεί τη μία, προσθέτει επιπλέον βοηθητικές συναρτήσεις για την καλύτερη υλοποίηση της λειτουργίας του αλγόριθμου και ορισμένες ιδιότητες που

χρησιμοποιούνται για να αποθηκεύουν τοπικά τις διάφορες παραμέτρους έτσι ώστε να μη χρειάζεται συνεχώς η ανάκτησή τους από τα προφίλ των διεργασιών και του workflow.

Η κλάση *Node* αποτελεί μια βοηθητική κλάση που αποθηκεύει στις ιδιότητές της τις παραμέτρους κόστους και αξιοπιστίας κάθε υποψήφιας διεργασίας, ώστε να μη γίνεται συνεχώς ανάκληση των δεδομένων αυτών από τα αντικείμενα τύπου διεργασίας. Επίσης χρησιμοποιείται για να προσθέτει επιπλέον ιδιότητες που χρησιμοποιεί ο αλγόριθμος, όπως ο απλός λόγος κόστος/διαθεσιμότητα και ο τροποποιημένος λόγος κόστος/διαθεσιμότητα.

Η κλάση *Normalizer* αναπαριστά τη συνάρτηση κανονικοποίησης που χρησιμοποιεί ο αλγόριθμος. Η συνάρτηση δεν ενσωματώθηκε στη διεπαφή,, για να μειωθεί η πολυπλοκότητα του κώδικα και επίσης για λόγους αφαίρεσης, ώστε να υπάρχει η δυνατότητα στο μέλλον επιλογής της συνάρτησης κανονικοποίησης. Δηλαδή μπορεί να υπάρχουν πολλές κλάσεις, που καθεμία να υλοποιεί διαφορετική συνάρτηση κανονικοποίησης.

Το διάγραμμα των παραπάνω κλάσεων με τις βασικές μεθόδους και ιδιότητες παρουσιάζεται στο σχήμα της επόμενης σελίδας (Σχήμα 44).



Σχήμα 44: Διάγραμμα κλάσεων του μηχανισμού απεικόνισης ροών εργασίας

6.7 Σύγκριση Διαμεσολαβητή και Μηχανισμού Απεικόνισης

Τόσο ο διαμεσολαβητής (broker) όσο και ο μηχανισμός απεικόνισης δυναμικών ροών εργασίας θεωρούνται υπηρεσίες απόφασης καθώς επιλέγουν υπηρεσίες από ένα σύνολο υποψηφίων. Η διαφορά μεταξύ των δύο στοιχείων είναι ότι ο μηχανισμός απεικόνισης επιλέγει ένα σύνολο υποψηφίων από τις δεδομένες αφηρημένες διεργασίες μιας ροής εργασίας και επιχειρεί τη βελτιστοποίηση της αρχικής επιλογής βάσει των παραμέτρων ποιότητας και τους στόχους χρηστών. Ο διαμεσολαβητής βασίζει τις αποφάσεις επιλογής σχετικά με τη δυνατότητα των υπηρεσιών να παρουσιάσουν μια αποδεκτή σύμβαση παροχής υπηρεσιών SLA. Επομένως, τα κριτήρια για την επιλογή των ιδιαίτερων υπηρεσιών είναι διαφορετικά.

Η σελίδα αυτή είναι σκόπιμα λευκή

Με εξέταση των δεδομένων εισόδου και εξόδου των δύο στοιχείων φαίνεται άμεσα η ομοιότητά τους καθώς απαιτούν ένα σύνολο υποψήφιων υπηρεσιών και απαιτήσεων χρηστών αναφορικά με την ποιότητα υπηρεσίας. Δεδομένου ότι και τα δύο συστατικά χρησιμοποιούν τα διαφορετικά κριτήρια επιλογής, απαιτούνται διαφορετικά σύνολα πρόσθετων πληροφοριών για την εκτέλεση της επιλογής. Ο μηχανισμός απεικόνισης χρειάζεται τις πληροφορίες ποιότητας για τις υποψήφιες υπηρεσίες ενώ ο διαμεσολαβητής χρειάζεται τα πρότυπα συμβόλαια SLA.

Τέλος, υπάρχει και μια σημαντική διαφορά μεταξύ τους. Ο μηχανισμός απεικόνισης ροών εργασίας λειτουργεί για το σύνολο των ροών ενώ ο διαμεσολαβητής μπορεί μόνο να χειριστεί τους στόχους των συγκεκριμένων διεργασιών των ροών και όχι του συνόλου αυτών. Αν και ο διαμεσολαβητής θα μπορούσε να επεκταθεί ώστε να συγκεντρώνει όλους τους στόχους σε μια ροή εργασίας και να διαπραγματεύεται συμβόλαια SLAs για όλους τους στόχους, το αποτέλεσμα θα ήταν ακόμα διαφορετικό από αυτό που παρέχει ο μηχανισμός απεικόνισης. Ο διαμεσολαβητής κάθε φορά θα εξέταζε το κάθε επίπεδο χωριστά ενώ ο μηχανισμός απεικόνισης εκτελεί βελτιστοποίηση της ροής εργασίας (διαφορετική επιλογή των υποψηφίων βασισμένη στις παραμέτρους ποιότητας) τόσο σε κάθε επίπεδο / τύπο διεργασίας αυτής όσο και για το σύνολο της ροής.

Η σελίδα αυτή είναι σκόπιμα λευκή

7

Αξιολόγηση του

Προτεινόμενου Μηχανισμού

Στο συγκεκριμένο κεφάλαιο γίνεται αξιολόγηση της λειτουργίας και της απόδοσης του μηχανισμού απεικόνισης δυναμικών ροών εργασίας και επιλογής υπηρεσιών καθώς και της αποτελεσματικότητάς του με χρήση ενός σεναρίου σε περιβάλλον πλέγματος για την εφαρμογή επεξεργασίας τρισδιάστατων εικόνων (3D Rendering). Επίσης παρουσιάζεται το σύστημα δοκιμών και τρεις (3) προσομοιώσεις με τα αντίστοιχα αποτελέσματα.

7.1 Εφαρμογή

Στη παρούσα παράγραφο παρουσιάζονται λεπτομέρειες για την υπηρεσία (service) που υλοποιεί τη λειτουργικότητα του μηχανισμού απεικόνισης ροών εργασίας η οποία είναι συμβατή με την αρχιτεκτονική OGSA (Open Grid Services Architecture) όπως αυτή περιγράφεται στην ενότητα 4.3.

Ο μηχανισμός αυτής της αρχιτεκτονικής σχεδιάστηκε βάσει των προδιαγραφών του πλαισίου WSRF (βλ. ενότητα 4.3.2.5) [36] και, επομένως, αποτελεί μια διαδικτυακή υπηρεσία (web service). Επίσης ο τρόπος υλοποίησης του μηχανισμού επιτρέπει τη χρήση του ως μια

αυτόνομη μονάδας προκειμένου να επιτευχθεί η βέλτιστη επιλογή υποψηφίων υπηρεσιών από τύπους υπηρεσίας ή από ολόκληρη τη ροή εργασίας σε σχέση με τις παραμέτρους ποιότητας που λαμβάνονται από τους παρόχους υπηρεσιών και από το χρήστη.

7.2 Σύστημα Δοκιμών

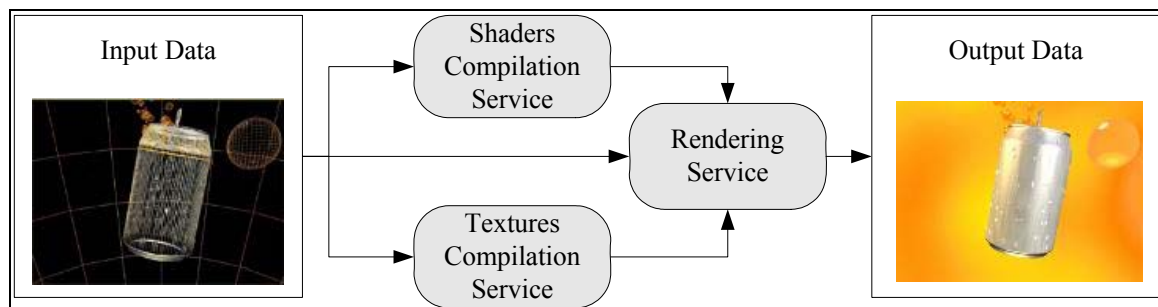
Για τους σκοπούς των προσομοιώσεων, χρησιμοποιήθηκε μια εφαρμογή επεξεργασίας τρισδιάστατων εικόνων σε περιβάλλον πλέγματος με χρήση το μεσολογισμικο (middleware) GRIA [13], [15]. Το GRIA είναι υλοποιημένο και λειτουργεί με διαδικτυακές υπηρεσίες, είναι συμβατό με OGSA και υλοποιήθηκε αρχικά στο πλαίσιο του έργου GRIA IST. Η υποδομή που χρησιμοποιήθηκε για τις προσομοιώσεις αποτελείται από μια φάρμα (cluster) οκτώ (8) υπολογιστών με τις ακόλουθες προδιαγραφές (Πίνακας 1):

CPU	Μνήμη	Χωρητικότητα δίσκου	Δίκτυο	Λειτουργικό σύστημα
x86 Intel P4 3GHz	RAM 512MB	Σκληρός δίσκος 80GB	Τοπικό δίκτυο (LAN0 1Gbps)	LINUX Fedora core 5

Πίνακας 1: Προδιαγραφές των υπολογιστών της φάρμας (cluster)

Η υπηρεσία επεξεργασίας τρισδιάστατων εικόνων βασίζεται σε έκδοση της εφαρμογής AIR [37]. Το GRIA που χρησιμοποιήθηκε ως μεσολογισμικό παρέχει την υπηρεσία επεξεργασίας συνδέοντας αρκετές ενδιαμέσες υπηρεσίες, όπως η συλλογή δεδομένων, ο έλεγχος της εγκυρότητας, η υπηρεσία έναρξης επεξεργασίας και η παράδοση του αποτελέσματος. Επίσης, προσφέρει τις υπηρεσίες *Shaders Compilation* και *Textures Compilation* που είναι προαπαιτούμενες για την υπηρεσία επεξεργασίας εικόνας (Rendering). Η επεξεργασία επιτυγχάνεται με χρήση περιγραφητών που εκφράζονται από τη μορφή RIB (RenderMan Interface Bytestream) [38] που είναι πλήρης προδιαγραφή της απαιτούμενης διασύνδεσης μεταξύ των εφαρμογών προσομοίωσης και επεξεργασίας. Ειδικότερα, η διαδικασία

επεξεργασίας επεκτείνεται με προσθήκη ενός ενδιάμεσου βήματος, όπου δημιουργείται ένα αρχείο RIB που περιγράφει την προς επεξεργασία σκηνή / εικόνα. Τα αρχεία σε μορφή RIB είναι σημαντικά μικρότερα σε μέγεθος και επομένως μεταφέρονται ευκολότερα μέσω δικτύου. Από την άλλη πλευρά, είναι πιο δύσκολα στη διαχείριση επειδή είναι πολυάριθμα και έχουν μεγάλο αριθμό διασυνδέσεων. Στη συνέχεια, εξειδικευμένες εφαρμογές επεξεργασίας (όπως το AIR) μπορούν να αναδομήσουν και να επεξεργαστούν την εικόνα με χρήση των ενδιάμεσων αρχείων RIB. Το αποτέλεσμα αυτής της διαδικασίας είναι η πλήρης επεξεργασμένη εικόνα. Η είσοδος της διαδικασίας περιλαμβάνει τα αρχεία εικόνων, τους Shaders και τα Textures ενώ η έξοδος είναι η τρισδιάστατη εικόνα όπως παρουσιάζεται στο παρακάτω σχήμα (σχήμα 45):



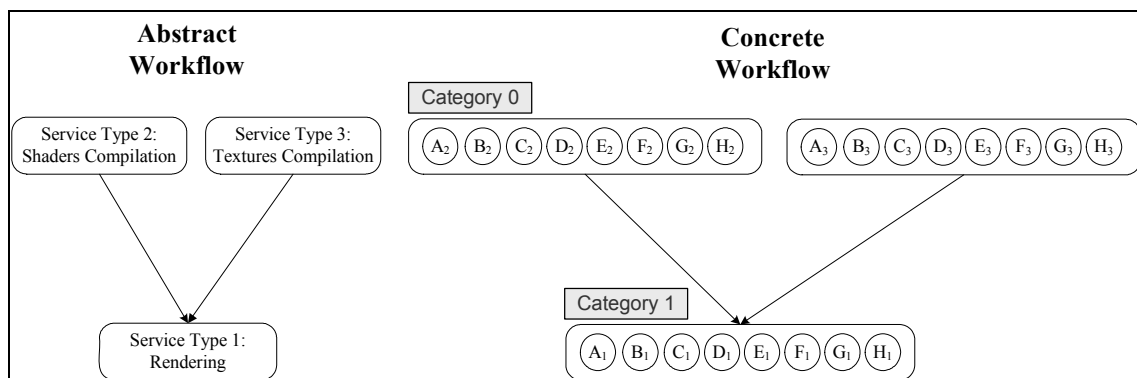
Σχήμα 45: Σενάριο παραγωγής τρισδιάστατης εικόνας

Το μεσολογισμικό GRIA προσφέρει αρκετές λειτουργίες στο χρήστη, όπως υποβολή εργασίας, παρακολούθηση εργασίας, επιλογή παρόχου υπηρεσίας και ορισμό προδιαγραφών QoS μέσω μιας διαδικασίας που καλείται "προσφορά" (tender) [39]. Ο χρήστης έχει πρόσβαση προς αυτές τις λειτουργίες μέσω της γραφικής διεπαφής χρήστη (GUI) του GRIA. Αυτή η διεπαφή GUI χρησιμοποιείται επίσης για να ληφθούν οι περιορισμοί / τα όρια που έχουν οριστεί από το χρήστη για τις παραμέτρους QoS καθώς και οι προτιμήσεις του χρήστη.

Η σελίδα αυτή είναι σκόπιμα λευκή

7.3 Αξιολόγηση

Σε αυτή την ενότητα, παρουσιάζονται τρεις (3) προσομοιώσεις που ολοκληρώθηκαν στο πλαίσιο του παραπάνω συστήματος δοκιμών καθώς και τα αντίστοιχα αποτελέσματά τους προς αξιολόγηση του μηχανισμού απεικόνισης ροών εργασίας. Οι προσομοιώσεις εκτελέστηκαν για δεδομένη ροή εργασίας εφαρμογής (application workflow) για την περίπτωση επεξεργασίας εικόνας που αποτελείται από τρεις (3) τύπους υπηρεσιών και οι οποίοι ταξινομούνται με βάση τη σημασία. Οι υπηρεσίες Shaders και Textures Compilation είναι προαπαιτούμενες για την υπηρεσία επεξεργασίας (Rendering). Επίσης, εφόσον οι υπηρεσίες Shaders και Textures Compilation μπορούν να εκτελεστούν παράλληλα, ανήκουν στην ίδια κατηγορία (όπως αναφέρεται στην ενότητα 6.4). Καθένας από τους παραπάνω τύπους υπηρεσιών περιλαμβάνει οκτώ (8) υποστάσεις υπηρεσιών – τους κόμβους της φάρμας (cluster) – που διαμορφώθηκαν ώστε να παρέχουν διαφορετικές παραμέτρους QoS για τους σκοπούς των πειραμάτων. Τα παραπάνω παρουσιάζονται στο σχήμα που ακολουθεί (Σχήμα 46):



Σχήμα 46: Αφηρημένη ροή εργασίας και αρχικοί τύποι & υποστάσεις υπηρεσιών

Στον πίνακα της επόμενης σελίδας (Πίνακας 2) δίνονται οι τιμές που ελήφθησαν για τις παραμέτρους ποιότητας (διαθεσιμότητα, κόστος και χρόνος) για όλες τις υποστάσεις υπηρεσιών.

Η σελίδα αυτή είναι σκόπιμα λευκή

Υπόσταση υπηρεσίας	Τιμή διαθεσιμότητας	Τιμή κόστους	Τιμή χρόνου
A ₁	96.50	180.32	300.25
B ₁	90.10	175.20	400.05
C ₁	87.30	165.66	442.43
D ₁	93.80	178.90	321.78
E ₁	92.40	177.75	320.65
F ₁	90.20	177.10	330.12
G ₁	89.50	170.24	370.94
H ₁	90.10	170.69	368.49
A ₂	90.68	70.35	3.27
B ₂	80.15	62.76	6.64
C ₂	97.66	79.39	9.15
D ₂	99.84	82.27	5.85
E ₂	70.15	60.19	7.53
F ₂	88.38	69.84	5.99
G ₂	95.03	74.44	3.12
H ₂	90.60	72.90	2.28
A ₃	98.01	110.44	4.67
B ₃	81.90	80.94	5.89
C ₃	95.83	99.73	7.40
D ₃	90.64	96.90	5.12
E ₃	89.95	87.19	6.10
F ₃	97.58	105.84	4.99
G ₃	95.10	100.34	4.48
H ₃	94.58	100.96	5.19

Πίνακας 2: Τιμές που ελήφθησαν αρχικά για τις υποστάσεις των υπηρεσιών

Με βάση αυτές τις τιμές, περιγράφονται στη συνέχεια τρεις αντιπροσωπευτικές προσομοιώσεις για την επιβεβαίωση και την αξιολόγηση του αλγορίθμου απεικόνισης δυναμικών ροών εργασίας και επιλογής υπηρεσιών και του σχήματος καθορισμού περισσότερων συγκεκριμένων ροών εργασίας.

7.3.1 1^η Προσομοίωση

Ο χρήστης ορίζει μέσω της διεπαφής τους ακόλουθους περιορισμούς για τις παραμέτρους QoS: (i) Ολική διαθεσιμότητα: Τουλάχιστον 80%, (ii) Ολικό κόστος: Όριο στις 330 μονάδες κόστους, (iii) Ολικός χρόνος εκτέλεσης: 200 μονάδες χρόνου. Επίσης, στην πρώτη

προσομοίωση, ο χρήστης δεν ορίζει βασικό παράγοντα (προτίμηση) και έτσι: $AvailabilitySlope = CostSlope = TimeSlope = 1$. Έτσι γίνεται προσομοίωση της περίπτωσης Βέλτιστης Επιλογής όπως περιγράφεται στην ενότητα 0.

Με βάση την περιγραφή του αλγορίθμου, η αρχική επιλογή είναι: υπόσταση C (C_1) από τον τύπο υπηρεσίας 1, υπόσταση B (B_2) από τον τύπο υπηρεσίας 2 και υπόσταση B (B_3) από τον τύπο υπηρεσίας 3.

Στον πίνακα της επόμενης σελίδας (Πίνακας 3), δίνονται οι επαναλήψεις του αλγορίθμου με βάση τα βήματα που περιγράφονται στην ενότητα 6.4 (η αύξηση των τιμών σημειώνεται με το σύμβολο "+" και η μείωση με το σύμβολο "-"). Για κάθε υπόσταση υπηρεσίας, υπάρχει δυνατή αντικατάσταση με βάση το δείκτη μετατροπής *ConvertedIndex* (όπως υπολογίζεται από την εξίσωση (17)) με τις αντίστοιχες διαφορές τιμών διαθεσιμότητας, κόστους και χρόνου. Για την πρώτη προσομοίωση και εφόσον ο τύπος υπηρεσίας 1 είναι υψηλής σημασίας (όπως δηλώνεται στη ροή εργασίας της εφαρμογής), στην πρώτη επανάληψη η αρχική υπόσταση υπηρεσίας C_1 αντικαθίσταται από την υπόσταση υπηρεσίας με το μικρότερο δείκτη (A_1). Στη συνέχεια, για τον τύπο υπηρεσίας 2, επιχειρείται η ίδια διαδικασία για τις υποστάσεις H_2 , G_2 και A_2 , οι οποίες έχουν μικρότερο δείκτη μετατροπής από τον αντίστοιχο δείκτη της επιλεγμένης υπόστασης υπηρεσίας (B_2). Ωστόσο, η αντικατάσταση αποτυγχάνει επειδή παραβιάζεται το όριο του συνολικού κόστους της ροής εργασίας. Το ίδιο ισχύει για τον τύπο υπηρεσίας 3. Τελικά, οι επόμενες βέλτιστες αντικαταστάσεις είναι η F_2 και η D_2 , οι οποίες αποτυγχάνουν επίσης λόγω του περιορισμού του κόστους. Αυτές οι αντικαταστάσεις επιχειρήθηκαν μετά την ολοκλήρωση των προσπαθειών για τον τύπο υπηρεσίας 3 καθώς, σύμφωνα με τα βήματα του αλγορίθμου, οι αντικαταστάσεις πραγματοποιούνται με βάση τις διαφορές τους ώστε να ορίζεται η σειρά αντικατάστασης των βέλτιστων υποστάσεων υπηρεσιών.

Επανάληψη	Αρχική υπόσταση υπηρεσίας	Πιθανή υπόσταση υπηρεσίας προς αντικατάσταση αρχικής	Διαφορά τιμής διαθεσιμότητας	Διαφορά τιμής κόστους	Διαφορά τιμής χρόνου	Αρχικός δείκτης (ConvertedIndex)	Πιθανός νέος δείκτης (ConvertedIndex)
1	C ₁	A ₁	+9.20	+14.66	-142.18	16862.87	561.05
2	B ₂	H ₂	+10.45	+10.14	-4.36	34.90	1.84
3	B ₂	G ₂	+14.88	+11.68	-3.52	34.90	3.53
4	B ₂	A ₂	+10.53	+7.59	-3.37	34.90	3.91
5	B ₃	G ₃	+13.20	+19.40	-1.41	154.68	4.73
6	B ₃	A ₃	+16.11	+29.50	-1.22	154.68	6.40
7	B ₃	F ₃	+15.68	+24.90	-0.90	154.68	9.14
8	B ₃	D ₃	+8.74	+15.96	-0.77	154.68	10.56
9	B ₃	H ₃	+12.68	+20.02	-0.70	154.68	11.49
10	B ₃	E ₃	+8.05	+6.25	+0.21	154.68	31.23
11	B ₂	D ₂	+19.69	+19.51	-0.79	34.90	22.91
12	B ₂	F ₂	+8.23	+7.08	-0.65	34.90	23.92

Πίνακας 3: Επαναλήψεις αλγορίθμου για την 1η προσομοίωση

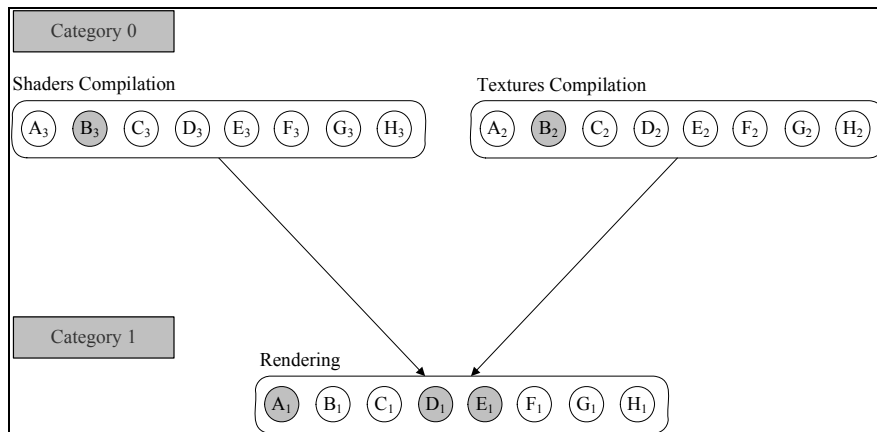
Μετά τις παραπάνω επαναλήψεις, δεν είναι δυνατές περαιτέρω αντικαταστάσεις και εφόσον δεν ικανοποιείται ο χρονικός περιορισμός λόγω της υπόστασης υπηρεσίας A_1 , αρχικοποιείται ο μηχανισμός χρονοπρογραμματισμού. Ο τύπος υπηρεσίας που επηρεάζει περισσότερο το χρόνο είναι ο τύπος 1, που ανήκει στην κατηγορία 1, ενώ οι τύποι υπηρεσιών 2 και 3 μπορούν να εκτελεστούν παράλληλα και, επομένως, ανήκουν στην ίδια κατηγορία (κατηγορία 0). Ο αρχικός χρόνος εκτέλεσης για την υπόσταση υπηρεσίας A_1 είναι 300,25 μονάδες χρόνου, ο συνολικός χρόνος εκτέλεσης της ροής εργασίας είναι 306,89 (ο χρόνος που χρειάζεται για να εκτελεστεί η A_1 συν ο χρόνος που απαιτείται για την εκτέλεση της B_2) και το συνολικό κόστος της ροής εργασίας είναι 324,02 μονάδες κόστους. Στον πίνακα της επόμενης σελίδας (Πίνακας 4) δίνονται οι επαναλήψεις του αλγορίθμου εντός του μηχανισμού χρονοπρογραμματισμού, όπου κάθε φορά επιλέγεται μια πρόσθετη υπόσταση υπηρεσίας με βάση το μικρότερο χρόνο εκτέλεσης.

Επανάληψη	Αρχική υπόσταση υπηρεσίας	Πιθανή πρόσθετη υπόσταση υπηρεσίας	Αρχικός χρόνος κρίσιμης υπηρεσίας	Νέος χρόνος κρίσιμης υπηρεσίας	Νέο κόστος ροής εργασίας
13	A_1	E_1	300.25	163.53	326.91
14	A_1	D_1	163.53	111.55	329.85

Πίνακας 4: Επαναλήψεις αλγορίθμου στο μηχανισμό χρονοπρογραμματισμού για την 1^η προσομοίωση

Οι επόμενες επαναλήψεις του αλγορίθμου δεν καταλήγουν σε νέες προσθήκες είτε επειδή παραβιάζεται ο περιορισμός κόστους ή επειδή αυξάνεται ο χρόνος εκτέλεσης με την προσθήκη νέας υπόστασης υπηρεσίας.

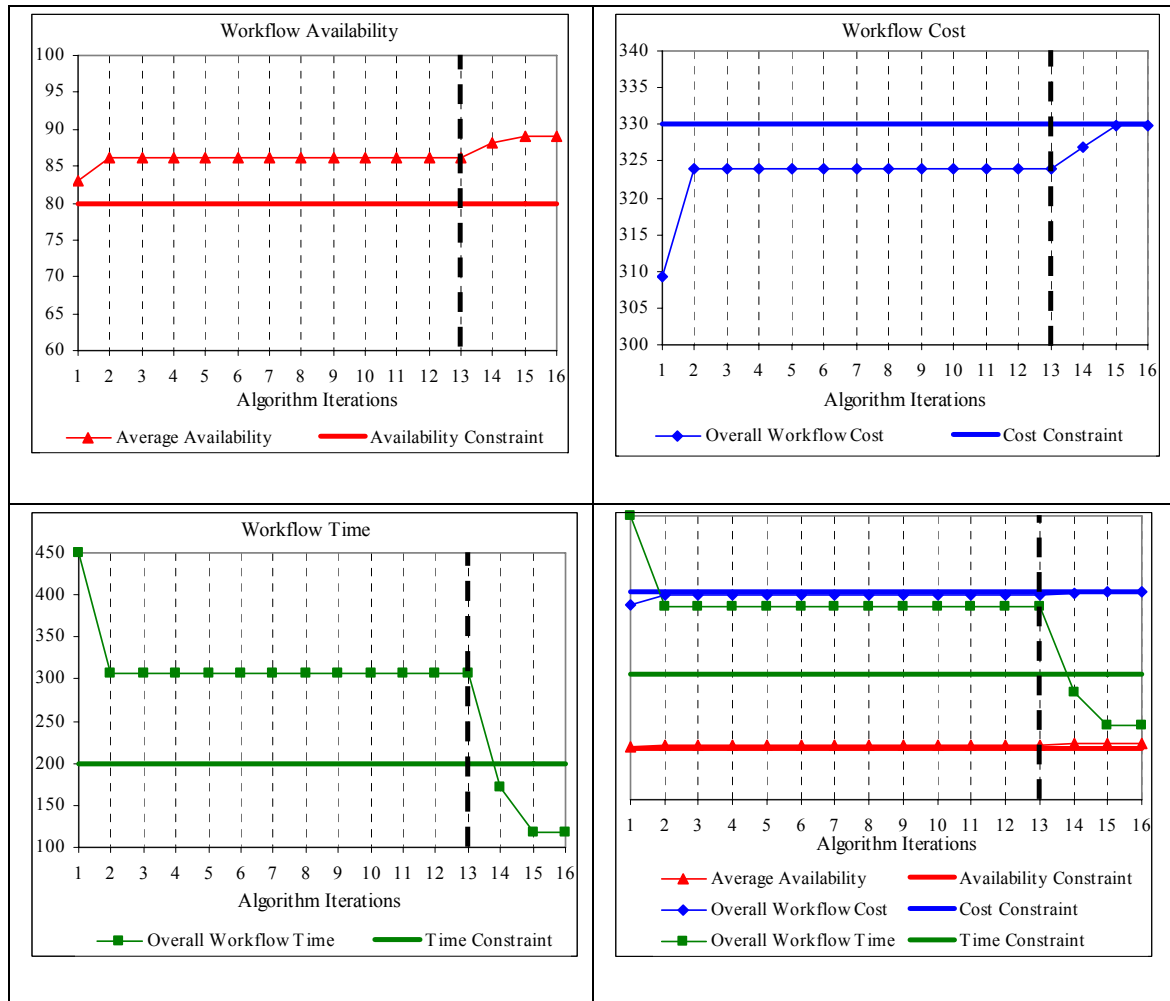
Στο σχήμα της επόμενης σελίδας (Σχήμα 47) παρουσιάζεται η τελική συγκεκριμένη ροή εργασίας (concrete workflow) για την 1^η Προσομοίωση.



Σχήμα 47: Συγκεκριμένη ροή εργασίας για την 1^η προσομοίωση

Επίσης, στο παρακάτω σχήμα (Σχήμα 48), δίνεται μια γραφική απεικόνιση των επαναλήψεων του αλγορίθμου για αυτή την προσομοίωση. Δείχνει τη μεταβολή της διαθεσιμότητας, του κόστους και του χρόνου εκτέλεσης συναρτήσεων των επαναλήψεων του αλγορίθμου (οι "έντονες" οριζόντιες γραμμές δηλώνουν τους περιορισμούς του χρήστη και η "διακεκομμένη" κατακόρυφη γραμμή υποδεικνύει την έναρξη του μηχανισμού χρονοπρογραμματισμού). Εφόσον ο χρήστης δεν έχει ορίσει καμία προτίμηση (κύριος παράγοντας επιλογής), το σχήμα δείχνει ότι ο μηχανισμός πραγματοποίησε τις επιλογές με βάση τους περιορισμούς του χρήστη (όρια στις παραμέτρους ποιότητας) ενώ στην επανάληψη 13 αρχικοποιείται ο μηχανισμός χρονοπρογραμματισμού προκειμένου να ικανοποιηθεί ο χρονικός περιορισμός.

Η σελίδα αυτή είναι σκόπιμα λευκή



Σχήμα 48: Μεταβολή τιμών διαθεσιμότητας, κόστους και χρόνου με βάση τις επιλογές που πραγματοποιούνται κατά τις επαναλήψεις του αλγορίθμου για την 1^η προσομοίωση

Η σελίδα αυτή είναι σκόπιμα λευκή

7.3.2 2^η Προσομοίωση

Σε αυτή την προσομοίωση, ο χρήστης ορίζει μέσω της διεπαφής GUI τους ακόλουθους περιορισμούς για τις παραμέτρους ποιότητας: (i) Ολική διαθεσιμότητα: Τουλάχιστον 80%, (ii) Ολικό κόστος: Όριο στις 350 μονάδες κόστους, (iii) Ολικός χρόνος εκτέλεσης: 200 μονάδες χρόνου. Επίσης, ο χρήστης ορίζει το χρόνο εκτέλεσης ως τον κύριο παράγοντα (προτίμηση) για τη διαδικασία επιλογής και έτσι: $AvailabilitySlope=CostSlope = 0$ και $TimeSlope=3$. Έτσι γίνεται προσομοίωση της περίπτωση Βελτιστοποίησης Χρόνου όπως περιγράφεται στην ενότητα 0.

Με βάση την περιγραφή του αλγορίθμου, η αρχική επιλογή είναι η υπόσταση C (C_1) από τον τύπο υπηρεσίας 1, η υπόσταση B (B_2) από τον τύπο υπηρεσίας 2 και η υπόσταση B (B_3) από τον τύπο υπηρεσίας 3.

Όμοια με την πρώτη προσομοίωση, στον πίνακα της επόμενης σελίδας (Πίνακας 5) δίνονται οι επαναλήψεις του αλγορίθμου. Εφόσον ο τύπος υπηρεσίας 1 είναι υψηλής σημασίας (όπως δηλώνεται στη ροή εργασίας της εφαρμογής), στην πρώτη επανάληψη η αρχική υπόσταση υπηρεσίας C_1 αντικαθίσταται από την υπόσταση υπηρεσίας με το μικρότερο δείκτη (A_1). Η ίδια διαδικασία ολοκληρώνεται επιτυχώς για τον τύπο υπηρεσίας 2 (αντικατάσταση της B_2 με την H_2) ενώ, για τον τύπο υπηρεσίας 3, η αντικατάσταση αποτυγχάνει λόγω παραβίασης του ορίου συνολικού κόστους της ροής εργασίας.

Επανάληψη	Αρχική υπόσταση υπηρεσίας	Πιθανή υπόσταση υπηρεσίας προς αντικατάσταση αρχικής	Διαφορά τιμής διαθεσιμότητας	Διαφορά τιμής κόστους	Διαφορά τιμής χρόνου	Αρχικός δείκτης (ConvertedIndex)	Πιθανός νέος δείκτης (ConvertedIndex)
1	C ₁	A ₁	+9.20	+14.66	-142.18	16862.87	561.05
2	B ₂	H ₂	+10.45	+10.14	-4.36	34.90	1.84
3	B ₃	G ₃	+13.20	+19.40	-1.41	154.68	4.73
4	B ₃	A ₃	+16.11	+29.50	-1.22	154.68	6.40
5	B ₃	F ₃	+15.68	+24.90	-0.90	154.68	9.14
6	B ₃	D ₃	+12.68	+20.02	-0.70	154.68	10.56
7	B ₃	H ₃	+8.74	+15.96	-0.77	154.68	11.49
8	B ₃	E ₃	+8.05	+6.25	+0.21	154.68	31.23

Πίνακας 5: Επαναλήψεις αλγορίθμου για τη 2^η προσομοίωση

Μετά τις παραπάνω επαναλήψεις, δεν είναι δυνατές περαιτέρω αντικαταστάσεις και, εφόσον δεν ικανοποιείται ο χρονικός περιορισμός λόγω της υπόστασης υπηρεσίας A_1 , εκκινεί ο μηχανισμός χρονοπρογραμματισμού. Ο τύπος υπηρεσίας που επηρεάζει περισσότερο το χρόνο είναι ο τύπος 1, που ανήκει στην κατηγορία 1, ενώ οι τύποι υπηρεσιών 2 και 3 μπορούν να εκτελεστούν παράλληλα και, επομένως, ανήκουν στην ίδια κατηγορία (κατηγορία 0). Ο αρχικός χρόνος εκτέλεσης για την υπόσταση υπηρεσίας A_1 είναι 300,25 μονάδες χρόνου, ο συνολικός χρόνος εκτέλεσης της ροής εργασίας είναι 306,89 (ο χρόνος που χρειάζεται για να εκτελεστεί η A_1 συν ο χρόνος που απαιτείται για την εκτέλεση της B_2) και το συνολικό κόστος της ροής εργασίας είναι 334,16 μονάδες κόστους. Στον πίνακα της επόμενης σελίδας (Πίνακας 6) δίνονται οι επαναλήψεις του αλγορίθμου εντός του μηχανισμού χρονοπρογραμματισμού, όπου κάθε φορά επιλέγεται μια πρόσθετη υπόσταση υπηρεσίας με βάση το μικρότερο χρόνο εκτέλεσης.

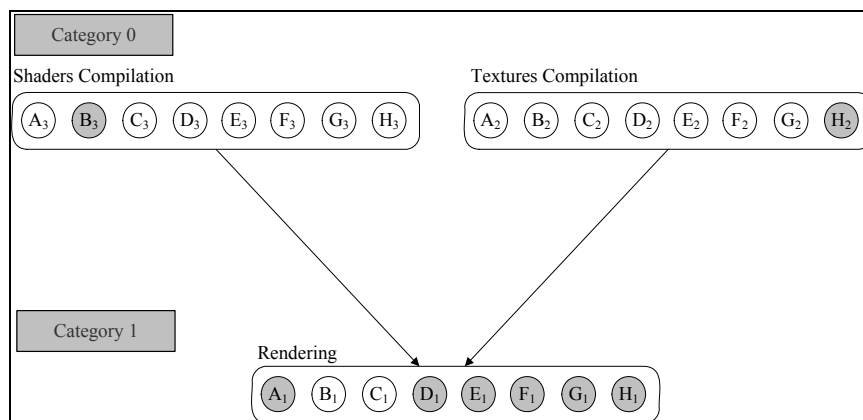
Επανάληψη	Αρχική υπόσταση υπηρεσίας	Πιθανή πρόσθετη υπόσταση υπηρεσίας	Αρχικός χρόνος κρίσιμης υπηρεσίας	Νέος χρόνος κρίσιμης υπηρεσίας	Νέο κόστος ροής εργασίας
9	A_1	E_1	300.25	163.53	336.46
10	A_1	D_1	163.53	111.55	340.00
11	A_1	F_1	111.55	87.48	343.07
12	A_1	H_1	87.48	79.59	344.95
13	A_1	G_1	79.59	68.01	347.26
14	A_1	B_1	68.01	64.01	350.67
15	A_1	C_1	68.01	64.15	349.48
16	B_3	G_3	5.89	3.00	358.77
17	H_2	G_2	2.28	1.59	349.50

Πίνακας 6: Επαναλήψεις αλγορίθμου στο μηχανισμό χρονοπρογραμματισμού για τη 2^η προσομοίωση

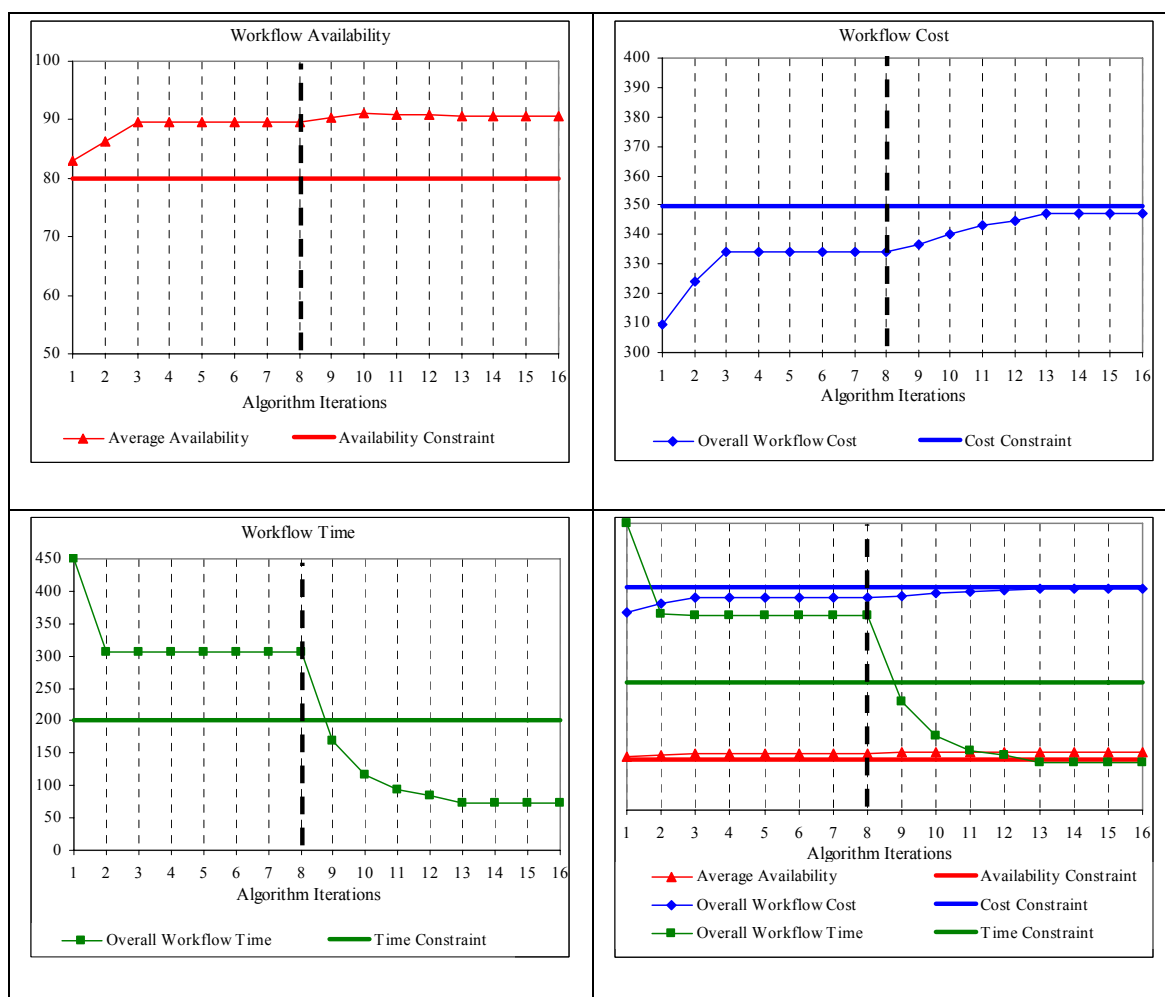
Οι πιθανές πρόσθετες υποστάσεις υπηρεσίας για τον τύπο υπηρεσίας 1 είναι: E_1 , D_1 , F_1 , H_1 , G_1 , B_1 και C_1 . Ωστόσο, η προσθήκη της B_1 αποτυγχάνει λόγω του ορίου συνολικού κόστους της ροής εργασίας και η προσθήκη της C_1 αποτυγχάνει επειδή αυξάνεται ο νέος χρόνος εκτέλεσης. Οι υπόλοιπες υποστάσεις υπηρεσίας μειώνουν το χρόνο εκτέλεσης του τύπου

υπηρεσίας 1 και, εφόσον δεν παραβιάζονται οι περιορισμοί του χρήστη, γίνονται οι αντίστοιχες προσθήκες. Επίσης, επιχειρείται βελτιστοποίηση χρόνου για τους τύπους υπηρεσιών 2 και 3 εφόσον αυτή η προσομοίωση δείχνει την περίπτωση ελαχιστοποίησης χρόνου που ισχύει για ολόκληρη τη ροή εργασίας. Ωστόσο, δεν είναι δυνατή η προσθήκη της υπόστασης υπηρεσίας G_2 επειδή η άλλη υπόσταση υπηρεσίας G_3 του τύπου υπηρεσίας που ανήκει στην ίδια κατηγορία δεν μπορεί να διαιρεθεί λόγω του ορίου για το συνολικό κόστος της ροής εργασίας.

Όμοια με την πρώτη προσομοίωση, στη σελίδα που ακολουθεί παρουσιάζεται η τελική συγκεκριμένη ροή εργασίας για τη 2^η προσομοίωση (Σχήμα 49) και η γραφική απεικόνιση των επαναλήψεων του αλγορίθμου (Σχήμα 50). Δείχνει τη μεταβολή της διαθεσιμότητας, του κόστους και του χρόνου εκτέλεσης συναρτήσει των επαναλήψεων του αλγορίθμου (οι "έντονες" οριζόντιες γραμμές δηλώνουν τους περιορισμούς του χρήστη και η "διακεκομμένη" κατακόρυφη γραμμή υποδεικνύει την έναρξη του μηχανισμού χρονοπρογραμματισμού). Εφόσον ο χρήστης ορίζει ως προτίμηση (κύριος παράγοντας επιλογής) τη βελτιστοποίηση χρόνου, το σχήμα δείχνει ότι ο μηχανισμός πραγματοποιεί τις επιλογές με βάση τους περιορισμούς του χρήστη (όρια στις παραμέτρους ποιότητας) ενώ στην επανάληψη 8 αρχικοποιείται ο μηχανισμός χρονοπρογραμματισμού προκειμένου να ελαχιστοποιηθεί ο συνολικός χρόνος εκτέλεσης της ροής εργασίας. Από την επανάληψη 8 έως την επανάληψη 16 δαπανιέται ο διαθέσιμος προϋπολογισμός (όριο κόστους) προκειμένου να ελαχιστοποιηθεί ο χρόνος. Για την αξιολόγηση του αλγορίθμου πραγματοποιήθηκαν πολλά πειράματα και τα παραπάνω είναι μόνο ενδεικτικά. Ο μηχανισμός καταλήγει στη συγκεκριμένη ροή εργασίας ακόμα κι αν μόνο ένας τύπος υπηρεσίας είναι είσοδος για την τελική απεικόνιση της ροής εργασίας.



Σχήμα 49: Συγκεκριμένη ροή εργασίας για τη 2^η προσομοίωση



Σχήμα 50: Μεταβολή τιμών διαθεσιμότητας, κόστους και χρόνου με βάση τις επιλογές που πραγματοποιούνται κατά τις επαναλήψεις του αλγορίθμου για τη 2^η προσομοίωση

Η σελίδα αυτή είναι σκόπιμα λευκή

7.3.3 3^η Προσομοίωση

Στη συγκεκριμένη προσομοίωση παρουσιάζονται τα αποτελέσματα από το συγχρονισμό των στοιχείων που περιγράφησαν στην ενότητα 6.5 αναφορικά με τον καθορισμό εναλλακτικών εκδόσεων συγκεκριμένων ροών εργασίας για εκτέλεση σε διαφορετικές χρονικές στιγμές.

Στον πίνακα της επόμενης σελίδας (πίνακας 7) αναγράφονται οι τιμές που ελήφθησαν αρχικά για τις παραμέτρους ποιότητας για τις υποστάσεις υπηρεσιών. Δεδομένου ότι η τιμή της παραμέτρου διαθεσιμότητας λαμβάνεται από τις υποστάσεις υπηρεσιών, οι τιμές χαρακτηρισμού που εμφανίζονται στον πίνακα υπολογίζονται με τη χρήση των εξισώσεων (18) και (19). Επιπλέον, ο πίνακας περιλαμβάνει τις τιμές παραμέτρων ποιότητας που αναφέρονται στην εκτέλεση μιας συγκεκριμένης ροής της δουλειάς σε διαφορετικό χρονικό πλαίσιο. Η λέξη "ιστορική" είναι στην παρένθεση δεδομένου ότι αυτό ισχύει μόνο εάν οι φορείς παροχής υπηρεσιών δε δημοσιεύουν (μέσω πρότυπων συμβολαίων SLAs) τις τιμές ποιότητας για διαφορετικά χρονικά πλαίσια σχετικά με την εκτέλεση της συγκεκριμένης ροής εργασίας. Σε αυτήν την περίπτωση, οι τιμές λαμβάνονται από τα ιστορικά στοιχεία που αποθηκεύονται στην τοπική βάση δεδομένων του τμήματος υποστήριξης αποφάσεων (βλ. ενότητα 6.5.1.2).

Σε αυτή την προσομοίωση, ο χρήστης καθορίζει τους ακόλουθους περιορισμούς για τις παραμέτρους ποιότητας: (i) Χαρακτηρισμός των φορέων παροχής υπηρεσιών: Τουλάχιστον 85,75%, (ii) Συνολικό Κόστος ροής εργασίας: Ανώτατο όριο 143.82 μονάδες κόστος, (iii) Συνολικός χρόνος εκτέλεσης ροής εργασίας: Ανώτατο όριο 170.24 χρονικές μονάδες. Επιπλέον, ο χρήστης καθορίζει τις προτιμήσεις που είναι μεταφρασμένες στους παράγοντες κλίσεων: $CostSlope = 1.13$, $TimeSlope = 0.85$, $CharacterizationSlope = 1.02$.

Βάσει των παραπάνω αρχικοποιείται το σχήμα καθορισμού περισσότερων συγκεκριμένων ροών εργασίας και τα αποτελέσματα της προσομοίωσης παρουσιάζονται ακολούθως.

Υπόσταση υπηρεσίας	Τιμή χαρακτηρισμού	Τιμή κόστους	Τιμή Χρόνου	(Ιστορική) Τιμή χαρακτηρισμού	(Ιστορική) Τιμή κόστους	(Ιστορική) Τιμή χρόνου
A ₁	93.19	90.77	321.11	96.12	88.85	312.14
B ₁	98.19	95.62	378.65	99.82	93.24	345.23
C ₁	80.92	73.33	482.25	85.34	68.56	410.24
D ₁	93.27	89.83	333.52	95.48	86.34	320.13
E ₁	88.82	92.37	347.19	90.93	88.74	332.55
F ₁	92.13	85.12	330.98	93.23	82.47	321.67
G ₁	85.04	62.43	459.32	88.86	61.98	432.43
H ₁	89.38	79.47	352.13	92.97	77.97	340.42
A ₂	93.99	32.47	5.93	96.03	28.12	5.64
B ₂	92.39	28.22	12.23	94.83	25.75	11.23
C ₂	74.24	12.13	15.34	83.43	11.89	13.78
D ₂	88.23	19.28	10.32	91.32	16.83	8.23
E ₂	65.53	23.27	8.12	82.23	21.23	7.69
F ₂	97.34	32.32	4.46	98.92	27.53	4.01
G ₂	94.34	19.83	18.43	96.75	16.25	16.24
H ₂	91.23	28.22	3.99	95.67	27.23	3.69
A ₃	92.42	47.89	52.12	95.99	44.67	50.76
B ₃	99.13	38.64	67.93	99.56	35.12	64.29
C ₃	95.85	32.89	72.33	97.46	29.23	69.79
D ₃	85.19	45.46	50.54	88.64	42.88	49.48
E ₃	84.23	39.08	64.88	88.15	37.23	62.89
F ₃	93.98	35.83	68.85	95.75	32.18	65.32
G ₃	92.29	41.11	68.63	94.29	38.67	66.63
H ₃	89.04	29.27	82.49	89.04	27.33	77.32

Πίνακας 7: Τιμές που ελήφθησαν αρχικά για τις υποστάσεις των υπηρεσιών

Σύμφωνα με τη λειτουργικότητα που περιγράφεται στην ενότητα 6.5.2, η αρχική απεικόνιση της ροής εργασίας οδηγεί σε μια αρχική επιλογή που είναι: υπόσταση υπηρεσίας H (H_1) από τον τύπο υπηρεσιών 1 - rendering, υπόσταση υπηρεσίας D (D_2) από τον τύπο υπηρεσιών 2 - shaders compilation - και υπόσταση υπηρεσίας H (H_3) από τον τύπο υπηρεσιών 3 – textures compilation.

Στον πίνακα της επόμενης σελίδας (Πίνακας 8), παρουσιάζουμε τις επαναλήψεις του αλγορίθμου (η αύξηση στις τιμές σημειώνεται με το σύμβολο "+" και η μείωση με το σύμβολο "-"). Για κάθε περίπτωση υπηρεσιών, υπάρχει μια πιθανή αντικατάσταση βασισμένη στο *ConvertedIndex* - όπως καθορίζεται στην εξίσωση (17), με τον αντίστοιχο χαρακτηρισμό, και τις τιμές των παραμέτρων κόστους και χρόνου. Καθώς ο τύπος υπηρεσίας 1 είναι υψηλής σημασίας (όπως δηλώνεται στη ροή εργασίας της εφαρμογής), στην πρώτη επανάληψη η αρχική υπόσταση υπηρεσίας C_1 αντικαθίσταται από την υπόσταση υπηρεσίας με το μικρότερο δείκτη (F_1). Κατόπιν, για τον τύπο υπηρεσίας 2, η ίδια διαδικασία επιχειρείται για τις υποστάσεις υπηρεσίας H_2 και το F_2 που έχουν χαμηλότερο *ConvertedIndex* από την επιλεγμένη υπόσταση υπηρεσίας (D_2). Εντούτοις, οι αντικαταστάσεις αποτυγχάνουν λόγω παραβίασης του συνολικού ορίου κόστους της ροής εργασίας. Η ίδια διαδικασία ολοκληρώνεται για τον τύπο υπηρεσίας 3 που οδηγεί στην αντικατάσταση του H_3 από το C_3 .

Μετά τις παραπάνω επαναλήψεις, δεν είναι δυνατές περαιτέρω αντικαταστάσεις και εφόσον δεν ικανοποιείται ο χρονικός περιορισμός λόγω της υπόστασης υπηρεσίας F_1 , αρχικοποιείται ο μηχανισμός χρονοπρογραμματισμού. Ο τύπος υπηρεσίας που επηρεάζει περισσότερο το χρόνο είναι ο τύπος 1, που ανήκει στην κατηγορία 1, ενώ οι τύποι υπηρεσιών 2 και 3 μπορούν να εκτελεστούν παράλληλα (επομένως ανήκουν στην ίδια κατηγορία: κατηγορία 0).

Επανάληψη	Αρχική υπόσταση υπηρεσίας	Πιθανή υπόσταση υπηρεσίας προς αντικατάσταση αρχικής	Διαφορά τιμής χαρακτηρισμού	Διαφορά τιμής κόστους	Διαφορά τιμής χρόνου	Αρχικός δείκτης (ConvertedIndex)	Πιθανός νέος δείκτης (ConvertedIndex)
1	H ₁	F ₁	+2.75	+5.65	-21.12	399.65	359.75
2	D ₂	H ₂	+3.00	+8.94	-6.33	2.35	1.32
3	D ₂	F ₂	+9.11	+13.04	-5.86	2.35	1.68
4	H ₃	C ₃	+13.20	+19.40	-1.41	45.64	24.92

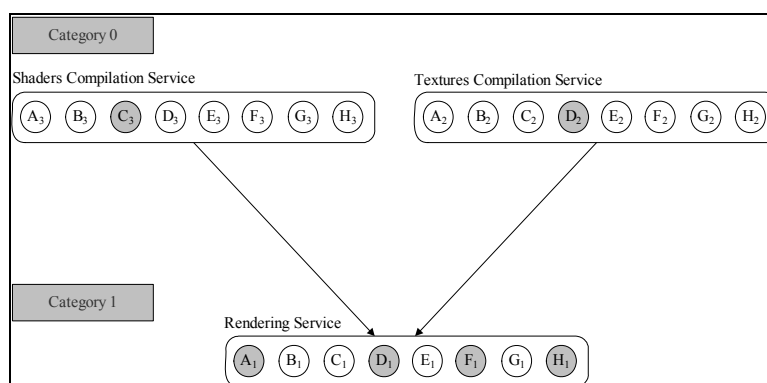
Πίνακας 8: Επαναλήψεις αλγορίθμου για την αρχική έκδοση συγκεκριμένης ροής εργασίας

Ο αρχικός χρόνος εκτέλεσης για την υπηρεσία F_1 είναι 330.98 μονάδες χρόνου, ο συνολικός χρόνος εκτέλεσης της ροής εργασίας είναι 403.31 (ο χρόνος που χρειάζεται για να εκτελεστεί η F_1 συν ο χρόνος που απαιτείται για την εκτέλεση της C_3) και το συνολικό κόστος της ροής εργασίας είναι 137.29 μονάδες κόστους. Στον παρακάτω πίνακα (Πίνακας 9) δίνονται οι επαναλήψεις του αλγορίθμου εντός του μηχανισμού χρονοπρογραμματισμού (επιλέγεται μια πρόσθετη υπηρεσία βάσει χρόνου εκτέλεσης).

Επανάληψη	Αρχική υπόσταση υπηρεσίας	Πιθανή πρόσθετη υπόσταση υπηρεσίας	Αρχικός χρόνος κρίσιμης υπηρεσίας	Νέος χρόνος κρίσιμης υπηρεσίας	Νέο κόστος ροής εργασίας
5	F_1	A_1	330.98	168.80	141.87
6	F_1	H_1	168.80	122.07	140.70
7	F_1	D_1	122.07	93.31	143.65

Πίνακας 9: Επαναλήψεις αλγορίθμου στο μηχανισμό χρονοπρογραμματισμού για την αρχική συγκεκριμένη ροή εργασίας

Στο παρακάτω σχήμα (Σχήμα 51) παρουσιάζεται η τελική συγκεκριμένη ροή εργασίας που θα εκτελεσθεί άμεσα και παρέχεται στο χρήστη ως αρχική έκδοση βασισμένη στους περιορισμούς ποιότητας και τις προτιμήσεις του.



Σχήμα 51: Αρχική έκδοση συγκεκριμένης ροής εργασίας

Όπως έχει ήδη αναφερθεί (βλ. ενότητα 6.5), η προστιθέμενη αξία της συνδυασμένης χρήσης των στοιχείων βρίσκεται στην παροχή διαφορετικών συγκεκριμένων ροών εργασίας που όχι μόνο καλύπτουν τους περιορισμούς και τις απαιτήσεις του χρήστη αλλά επιπλέον

Η σελίδα αυτή είναι σκόπιμα λευκή

μεγιστοποιούν το όφελός του από την άποψη της προσφερθείσας ποιότητας. Η συγκεκριμένη προσομοίωση συνεχίζει προκειμένου να παρασχεθούν οι επιπλέον εκδόσεις της συγκεκριμένης ροής εργασίας, οι οποίες θα εκτελεσθούν σε διαφορετικά χρονικά πλαίσια και όχι άμεσα. Η πρώτη έκδοση συγκεκριμένης ροής εργασίας καθορίζεται κατά τέτοιο τρόπο ώστε να μεγιστοποιεί το χαρακτηρισμό των παρόχων που επιλέγονται και να ελαχιστοποιεί το συνολικό χρόνο εκτέλεσης της ροής εργασίας χωρίς την παραβίαση του περιορισμού κόστους που τίθεται από το χρήστη. Προκειμένου να επιτευχθεί αυτό, το στοιχείο διαχείρισης (βλ. ενότητα 6.5.1.1) θέτει τις εξής τιμές: $CostSlope=0.10$, $TimeSlope=1.45$, $CharacterizationSlope = 1.45$ και ζητά τον καθορισμό μιας νέας συγκεκριμένης ροής εργασίας από το μηχανισμό απεικόνισης. Επιπλέον, το στοιχείο διατήρησης ζητά τις τιμές των παραμέτρων ποιότητας των φορέων παροχής υπηρεσιών για τα διαφορετικά χρονικά πλαίσια από το τμήμα υποστήριξης αποφάσεων. Με βάση τις ανωτέρω πληροφορίες και όμοια με τον καθορισμό της αρχικής έκδοσης της συγκεκριμένης ροής εργασίας, οι ακόλουθοι πίνακες (Πίνακας 10 και Πίνακας 11) παρουσιάζουν τις επαναλήψεις του αλγορίθμου και το μηχανισμό χρονοπρογραμματισμού. Η αρχική επιλογή των περιπτώσεων υπηρεσιών είναι G_1 , G_2 και H_3 για τα οποία οι αλλαγές και οι προσθήκες εμφανίζονται στις επαναλήψεις του αλγορίθμου.

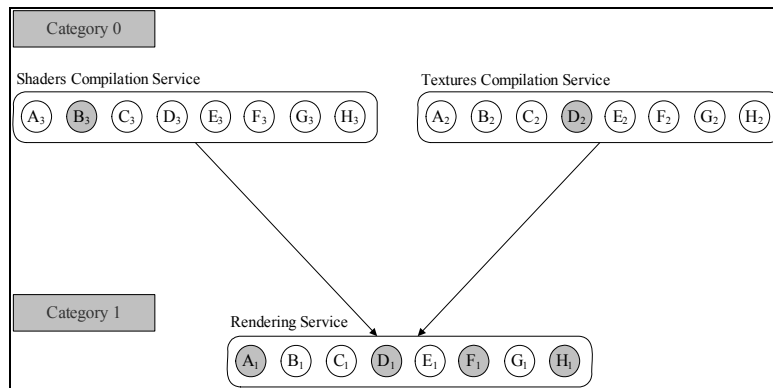
Επανάληψη	Αρχική υπόσταση υπηρεσίας	Πιθανή υπόσταση υπηρεσίας προς αντικατάσταση αρχικής	Διαφορά τιμής χαρακτηρισμού	Διαφορά τιμής κόστους	Διαφορά τιμής χρόνου	Αρχικός δείκτης (ConvertedIndex)	Πιθανός νέος δείκτης (ConvertedIndex)
1	G ₁	A ₁	+7.26	+26.87	-120.29	959.18	149.35
2	G ₂	F ₂	+2.17	+11.28	-12.23	4.76	0.47
3	G ₂	H ₂	-1.08	+10.98	-12.55	4.76	0.52
4	G ₂	A ₂	-0.72	+11.87	-10.60	4.76	0.99
5	G ₂	D ₂	-5.43	+0.58	-8.01	4.76	1.51
6	H ₃	A ₃	+6.95	+17.34	-26.56	89.02	14.01
7	H ₃	B ₃	+10.52	+7.79	-13.03	89.02	18.37

Πίνακας 10: Επαναλήψεις Αλγορίθμου για την 1^η εναλλακτική έκδοση συγκεκριμένης ροής εργασίας

Επανάληψη	Αρχική υπόσταση υπηρεσίας	Πιθανή πρόσθετη υπόσταση υπηρεσίας	Αρχικός χρόνος κρίσιμης υπηρεσίας	Νέος χρόνος κρίσιμης υπηρεσίας	Νέο κόστος ροής εργασίας
8	A ₁	D ₁	312.14	163.27	141.30
9	A ₁	F ₁	163.27	111.51	141.27
10	A ₁	H ₁	111.51	90.21	140.89

Πίνακας 11: Επαναλήψεις αλγορίθμου στο μηχανισμό χρονοπρογραμματισμού για την 1^η εναλλακτική έκδοση συγκεκριμένης ροής εργασίας

Με την ολοκλήρωση των παραπάνω επαναλήψεων καθορίζεται η πρώτη εναλλακτική συγκεκριμένη ροής εργασίας όπως παρουσιάζεται στο παρακάτω σχήμα (Σχήμα 52). Η συγκεκριμένη ροή μεγιστοποιεί το μέσο χαρακτηρισμό των παρόχων και ελαχιστοποιεί το συνολικό χρόνο εκτέλεσης της ροής με τη χρησιμοποίηση του διαθέσιμου προϋπολογισμού (όπως τίθεται με τον περιορισμό κόστος) για να επιλέξει τις αντίστοιχες υπηρεσίες.



Σχήμα 52: 1^η Εναλλακτική έκδοση συγκεκριμένης ροής εργασίας

Η δεύτερη εναλλακτική συγκεκριμένη ροή εργασίας καθορίζεται κατά τέτοιο τρόπο ώστε να ελαχιστοποιεί το συνολικό κόστος της ροής εργασίας χωρίς παραβίαση του μέσου χαρακτηρισμού και του συνολικού χρόνου εκτέλεσης της ροής. Προκειμένου να επιτευχθεί αυτό και παρόμοιος με τον ορισμό της πρώτης εναλλακτικής ροής, το στοιχείο διαχείρισης καθορίζει τα εξής: $CostSlope=2.80$, $TimeSlope=0.10$, $CharacterizationSlope=0.10$ και ζητά τον καθορισμό μιας νέας συγκεκριμένης ροής εργασίας από το μηχανισμό απεικόνισης ζητώντας παράλληλα τις τιμές των παραμέτρων ποιότητας των φορέων παροχής υπηρεσιών για διαφορετικά χρονικά πλαίσια από το τμήμα υποστήριξης αποφάσεων. Με βάση τις ανωτέρω πληροφορίες και παρόμοιος με τον καθορισμό της πρώτης εναλλακτικής συγκεκριμένης ροής εργασίας, οι ακόλουθοι πίνακες (Πίνακας 12 και Πίνακας 13) παρουσιάζουν τις επαναλήψεις του αλγορίθμου και το μηχανισμό χρονοπρογραμματισμού. Η αρχική επιλογή των περιπτώσεων υπηρεσιών είναι G₁, G₂ και H₃ για τα οποία οι αλλαγές και οι προσθήκες εμφανίζονται στις επαναλήψεις του αλγορίθμου.

Η σελίδα αυτή είναι σκόπιμα λευκή

Επανάληψη	Αρχική υπόσταση υπηρεσίας	Πιθανή υπόσταση υπηρεσίας προς αντικατάσταση αρχικής	Διαφορά τιμής χαρακτηρισμού	Διαφορά τιμής κόστους	Διαφορά τιμής χρόνου	Αρχικός δείκτης (ConvertedIndex)	Πιθανός νέος δείκτης (ConvertedIndex)
1	G ₂	D ₂	-5.43	+0.58	-8.01	5.86	3.49

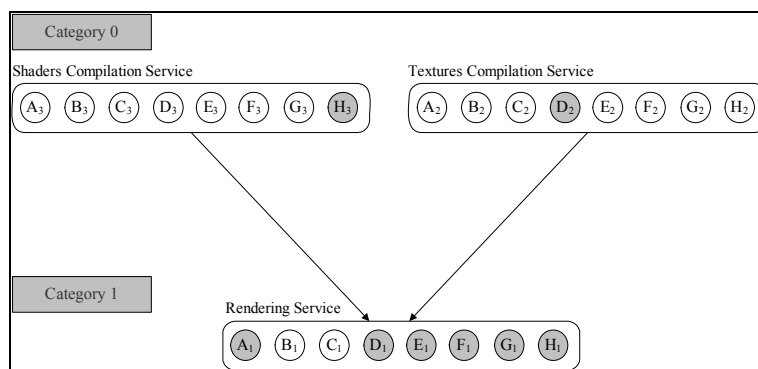
Πίνακας 12: Επαναλήψεις αλγορίθμου για την 2^η εναλλακτική έκδοση συγκεκριμένης ροής εργασίας

Επανάληψη	Αρχική υπόσταση υπηρεσίας	Πιθανή πρόσθετη υπόσταση υπηρεσίας	Αρχικός χρόνος κρίσιμης υπηρεσίας	Νέος χρόνος κρίσιμης υπηρεσίας	Νέο κόστος ροής εργασίας
2	G ₁	H ₁	432.43	220.54	115.53
3	G ₁	F ₁	220.54	149.91	121.27
4	G ₁	D ₁	149.91	114.59	125.98
5	G ₁	A ₁	149.91	93.41	130.04
6	G ₁	E ₁	93.41	79.28	133.32

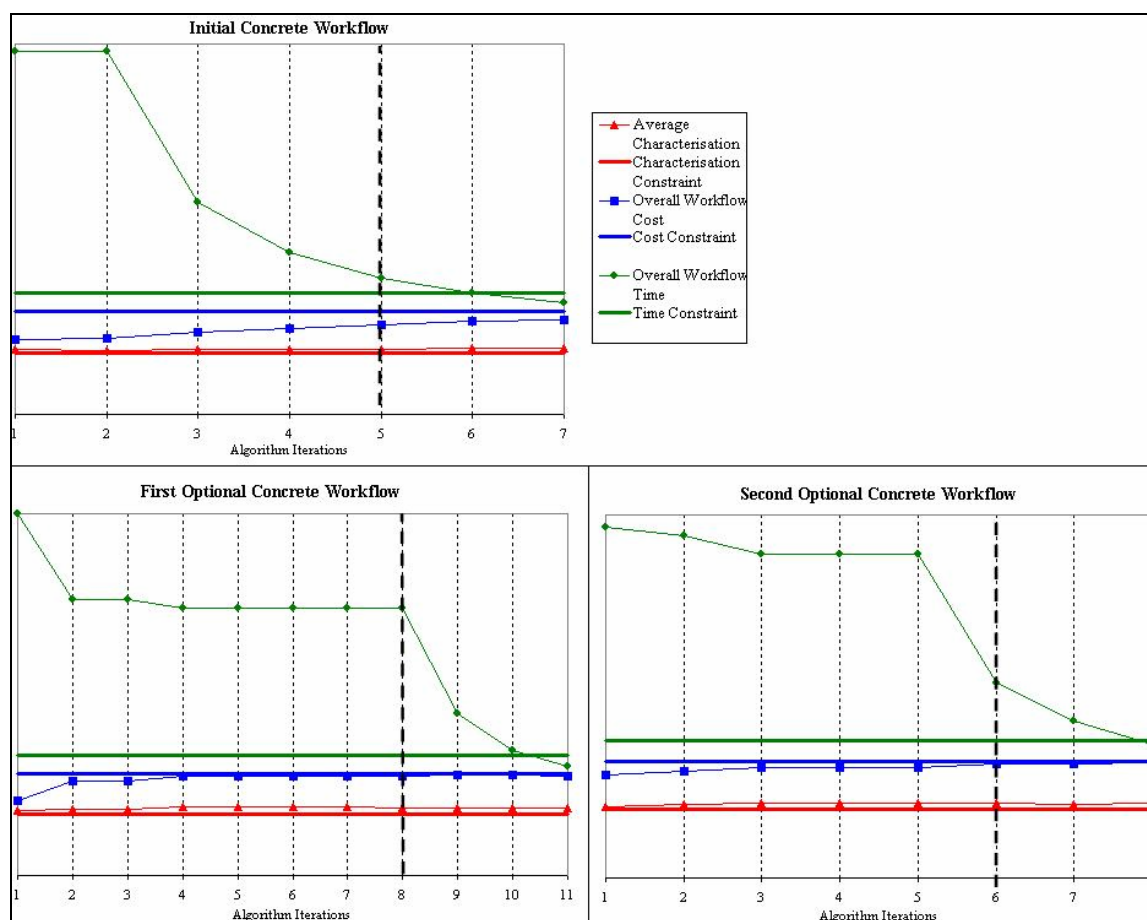
Πίνακας 13: Επαναλήψεις αλγορίθμου στο μηχανισμό χρονοπρογραμματισμού για τη 2^η εναλλακτική ροή εργασίας

Με την ολοκλήρωση των παραπάνω επαναλήψεων καθορίζεται η δεύτερη εναλλακτική συγκεκριμένη ροής εργασίας όπως παρουσιάζεται στο σχήμα της επόμενης σελίδας (Σχήμα 53). Αυτή η δεύτερη προαιρετική συγκεκριμένη ροή της δουλειάς συναντά τους περιορισμούς που τίθενται από το χρήστη σχετικά με το μέσο χαρακτηρισμό των παρόχων και το συνολικό χρόνο εκτέλεσης της ροής εργασίας αλλά ελαχιστοποιεί το γενικό κόστος αυτής.

Επίσης, στο επόμενο σχήμα (Σχήμα 54), δίνεται μια γραφική απεικόνιση των επαναλήψεων του αλγορίθμου για αυτή την προσομοίωση. Δείχνει τη μεταβολή του χαρακτηρισμού των παρόχων, του κόστους και του χρόνου εκτέλεσης συναρτήσει των επαναλήψεων του αλγορίθμου (οι "έντονες" οριζόντιες γραμμές δηλώνουν τους περιορισμούς του χρήστη και η "διακεκομμένη" κατακόρυφη γραμμή υποδεικνύει την έναρξη του μηχανισμού χρονοπρογραμματισμού).



Σχήμα 53: 2^η Εναλλακτική έκδοση συγκεκριμένης ροής εργασίας



Σχήμα 54: Μεταβολή τιμών χαρακτηρισμού, κόστους και χρόνου με βάση τις επιλογές που πραγματοποιούνται κατά τις επαναλήψεις του αλγορίθμου για τις τρεις εκδόσεις συγκεκριμένων ροών εργασίας

Η σελίδα αυτή είναι σκόπιμα λευκή

Τέλος, στον ακόλουθο πίνακα (Πίνακας 14) παρουσιάζεται μια περίληψη των παρεχόμενων ροών εργασίας, η αρχική - όπως ζητείται από το χρήστη για άμεση εκτέλεση και οι δύο (2) εναλλακτικές. Αυτή η περίληψη δείχνει ότι με τη χρησιμοποίηση του διαθέσιμου προϋπολογισμού (που καθορίζεται μέσω του περιορισμού κόστους του χρήστη) σε ένα διαφορετικό χρονικό πλαίσιο, ο συνολικός χρόνος της ροής εργασίας μειώνεται ενώ ο μέσος χαρακτηρισμός των επιλεγμένων φορέων παροχής υπηρεσιών αυξάνεται σε σύγκριση με τη συγκεκριμένη ροή εργασίας που θα εκτελούταν άμεσα. Εκτός αυτού, το κόστος ελαχιστοποιείται στην περίπτωση της εκτέλεσης σε ένα διαφορετικό χρονικό πλαίσιο ενώ οι περιορισμοί των άλλων δύο παραμέτρων δεν παραβιάζονται.

Επανάληψη	Συνολικό κόστος ροής εργασίας (μονάδες κόστους)	Ποσοστό Χαρακτηρισμού παρόχων (%)	Συνολικός χρόνος εκτέλεσης ροής εργασίας (μονάδες χρόνου)
Περιορισμοί Χρήστη	Μέγιστο 143.82	Όριο: Ελάχιστο 85.75	Όριο: Μέγιστο 170.24
Αρχική Συγκεκριμένη Ροή Εργασίας	143.64	92.01	165.64
1 ^η Εναλλακτική Συγκεκριμένη Ροή Εργασίας	140.89	94.78	154.50
2 ^η Εναλλακτική Συγκεκριμένη Ροή Εργασίας	133.32	92.24	156.59

Πίνακας 14: Σύνοψη επιπέδων προσφερόμενης ποιότητας υπηρεσίας για την αρχική και τις εναλλακτικές ροές εργασίας

Η σελίδα αυτή είναι σκόπιμα λευκή

8

Σύνοψη Κυριότερων

Συνεισφορών - Ανοιχτά

Ερευνητικά Θέματα

Στο συγκεκριμένο κεφάλαιο περιλαμβάνεται η σύνοψη της διατριβής και τα συμπεράσματα που εξήχθησαν κατά την εκπόνησή της καθώς και η κύρια συνεισφορά της στο χώρο της τεχνολογίας πλέγματος και συζητούνται θέματα μελλοντικής εργασίας και επέκτασης των ερευνητικών αποτελεσμάτων.

8.1 Σύνοψη Διατριβής και Κυριότερων Συνεισφορών

Στην παρούσα διδακτορική διατριβή παρουσιάστηκαν γενικά στοιχεία για τα περιβάλλον πλέγματος, καθώς και μια σειρά καινοτόμων μελετών και αναλύσεων για την Ποιότητα Υπηρεσιών και τις δυναμικές Ροές Εργασίας. Βάσει αυτών παρουσιάστηκε ένας νέος μηχανισμός απεικόνισης ροών εργασίας και επιλογής υπηρεσιών για διαχείριση και παροχή ποιότητας υπηρεσίας.

Ο προαναφερθέν μηχανισμός συμβάλλει στις προσπάθειες που γίνονται για να έχει το περιβάλλον πλέγματος δυναμική δυνατότητα παροχής ποιότητας υπηρεσιών. Στόχος του

μηχανισμού είναι ο ορισμός μιας συγκεκριμένης ροής εργασίας (concrete workflow) με βάση διάφορες παραμέτρους που δηλώνονται από το χρήστη και από τον πάροχο της υπηρεσίας, ενώ ταυτόχρονα λαμβάνονται υπόψη για την επιλογή των υπηρεσιών και οι προτιμήσεις του χρήστη (εκφράζονται ως βάρη επί των παραμέτρων). Για την επιλογή των υπηρεσιών χρησιμοποιείται ο καινοτόμος αλγόριθμος που αναπτύχθηκε και ο οποίος περιλαμβάνει και ένα σχήμα χρονοπρογραμματισμού για την επίτευξη χρονικών ορίων για το σύνολο της ροής εργασίας.

Καθώς η απεικόνιση των δυναμικών ροών εργασίας στοχεύει στη διασφάλιση της παρεχόμενης ποιότητας υπηρεσιών, στη διατριβή έγινε ανάλυση της έννοιας της ποιότητας υπηρεσίας, των παραμέτρων που την εκφράζουν και του τρόπου υπολογισμού αυτών καθώς και των κατηγοριών απαιτήσεων ποιότητας. Επιπλέον, παρουσιάζεται ένα σχήμα δρομολόγησης με στόχο την παροχή ποιότητας υπηρεσιών για συγκεκριμένες εργασίες.

Στη διατριβή παρουσιάζονται επίσης και αναλύονται συγκεκριμένα αρχιτεκτονικά θέματα όπως η Αντικειμενοστραφής Αρχιτεκτονική (SOA) και η Αρχιτεκτονική Ανοιχτών Υπηρεσιών (OGSA) βάσει των οποίων σχεδιάστηκε η καινοτόμος Αρχιτεκτονική Πλέγματος Επόμενης Γενιάς. Η ανάλυση των ανωτέρω ολοκληρώνεται με έμφαση σε συγκεκριμένα στοιχεία της αρχιτεκτονικής που επιτρέπουν το δυναμικό χαρακτήρα των ροών εργασίας. Βάσει αυτού παρουσιάζεται ένα νέο μοντέλο, το Μοντέλο Εικονικής Υποδομής (Grid VIM), ένα μοντέλο που στοχεύει στην αφαίρεση των επιχειρησιακών διεργασιών και των αλληλεπιδράσεων με την υποδομή πλέγματος κάνοντας χρήση προσαρμοσμένων ροών εργασίας και δυναμική δέσμευση υπηρεσιών, η οποία περιλαμβάνει σημασιολογική εύρεση και επιλογή αυτών. Ο μηχανισμός απεικόνισης δυναμικών εργασίας αποτελεί μέρος του παραπάνω μοντέλου και λόγω αυτού έγινε περιγραφή των αλληλεπιδράσεων και της επικοινωνίας αυτού με άλλα στοιχεία του μοντέλου.

Επιπρόσθετα παρατέθηκαν αποτελέσματα προσομοίωσης που αποδεικνύουν και αξιολογούν τη λειτουργία και αποδοτικότητα του προτεινόμενου μηχανισμού και της ενσωμάτωσης των παραμέτρων ποιότητας υπηρεσίας στο μηχανισμό. Δεδομένων των αποτελεσμάτων προσομοίωσης, ο μηχανισμός ολοκληρώνει την επιλογή των υπηρεσιών λαμβάνοντας υπόψη τους περιορισμούς και τις προτιμήσεις του χρήστη, γεγονός που τεκμηριώνει την απόδοση του μηχανισμού και επιτρέπει την αποδοχή του σε οποιοδήποτε ετερογενές σύστημα και, ειδικότερα, σε συστήματα πλέγματος που επιδιώκουν την εισαγωγή της γνώσης ποιότητας υπηρεσίας στη διαδικασία απεικόνισης δυναμικών ροών εργασίας.

Συμπερασματικά, αντίθετα με άλλα ετερογενή συστήματα [40], τα πλέγματα δεν έχουν ακόμα υιοθετήσει έναν αποτελεσματικό μηχανισμό που θα διευκολύνει τη παροχή ποιότητας υπηρεσίας για το σύνολο των ροών εργασίας και όχι για επιμέρους διεργασίες αυτών. Προς αυτή την κατεύθυνση, παρουσιάστηκε η σημασία της απεικόνισης ροών εργασίας μιας εφαρμογής σε υπηρεσίες πλέγματος λαμβάνοντας υπόψη το επίπεδο της ποιότητας υπηρεσίας που προσφέρεται από τους παρόχους υπηρεσιών πλέγματος και το αντίστοιχο επίπεδο που απαιτείται από τους χρήστες. Τα παραπάνω επίπεδα ποιότητα εκφράζονται με τη μορφή παραμέτρων που ουσιαστικά καθορίζουν και «ποσοτικοποιούν» την προσφερόμενη ποιότητα. Τα παραπάνω αναμένεται να συμβάλλουν στην υιοθέτηση τέτοιων μηχανισμών καθώς ο ρόλος των ροών εργασίας κρίνεται ολοένα και πιο σημαντικός σε υποδομές πλέγματος επιχειρηματικού χαρακτήρα [102].

8.2 *Ανοιχτά Ερευνητικά Θέματα*

8.2.1 *Στρατηγικές Εξαρτήσεις στις Ροές Εργασίας*

Η μελέτη που εκπονήθηκε στην παρούσα διατριβή επιτρέπει τη συζήτηση για ποικίλα ζητήματα μελλοντικής έρευνας στην περιοχή της επιλογής υπηρεσιών και της διασφάλισης

της ποιότητας των υπηρεσιών που προσφέρονται σε υποδομές πλέγματος. Μια ενδιαφέρουσα πτυχή είναι να επεκταθεί η συγκεκριμένη μελέτη ώστε η επιλογή των υποστάσεων υπηρεσιών ανά τύπο υπηρεσίας να γίνεται λαμβάνοντας υπόψη τυχόν επιχειρηματικές και στρατηγικές εξαρτήσεις (συνεργασίες ή ανταγωνιστικές σχέσεις) μεταξύ των παρόχων των υπηρεσιών. Οι προαναφερθείσες εξαρτήσεις θα επηρεάζουν τη διαδικασία επιλογής καθώς το κόστος για κάθε υπηρεσίας θα διαφοροποιείται βάσει προηγούμενων ή / και επόμενων επιλογών υπηρεσιών από διαφορετικούς φορείς παροχής υπηρεσιών που δύναται να έχουν στρατηγικές και επιχειρηματικές σχέσεις. Η υλοποίηση του παραπάνω απαιτεί μια σειρά ενεργειών, κάποιες από τις οποίες αναφέρονται στη συνέχεια:

- Ενημέρωση του τρόπου (γλώσσες) περιγραφής των ροών εργασίας ώστε να περιλαμβάνουν την παραπάνω πληροφορία ως επιπλέον μεταδεδομένα (metadata) των διεργασιών που περιλαμβάνονται στις ροές εργασίας.
- Διατύπωση των ποσοτικών όρων για την έννοια των στρατηγικών εξαρτήσεων ώστε να μπορεί να μεταφράζεται σε μετρικές που χρησιμοποιούνται από μηχανισμούς απεικόνισης ροών εργασίας.
- Υλοποίηση μηχανισμών και αλγορίθμων που προβλέπουν την επίδραση των εξαρτήσεων και των επιχειρηματικών πολιτικών στη διαδικασία τιμολόγησης των υπηρεσιών (ως επιπλέον κριτήριο επιλογής).

8.2.2 Μηχανισμός Καθορισμού Βέλτιστου Τρόπου Εκτέλεσης Εργασιών

Ένα σημαντικό θέμα που τίθεται καθώς οι υποδομές πλέγματος βρίσκουν εφαρμογή στον επιχειρηματικό κόσμο αναφέρεται στην απόφαση για την εκτέλεση και ολοκλήρωση εργασιών σε περιβάλλοντα πλέγματος. Αναλυτικότερα, δεν είναι σαφές για πολλές επιχειρήσεις αν είναι «βέλτιστο» να αποστείλουν προς εκτέλεση (ή αποθήκευση) τα

δεδομένα τους σε υποδομές πλέγματος. Ο όρος «βέλτιστο» αναφέρεται σε μια σειρά παραμέτρων στοιχείων όπως τα ακόλουθα:

- Χρόνος μεταφοράς δεδομένων σε πόρους πλέγματος, χρόνος επεξεργασίας και χρόνος ανάκτησης δεδομένων σε σύγκριση με το χρόνο εκτέλεσης εντός μιας επιχείρησης.
- Κόστος υπηρεσίας πλέγματος σε σύγκριση με το κόστος αγοράς, συντήρησης και υποδομής εντός μιας επιχείρησης.
- Διαθεσιμότητα και πρόσβαση από πολλαπλά σημεία στα δεδομένα σε σύγκριση με τη διαθεσιμότητα και τους τρόπους πρόσβασης σε μια επιχείρηση.
- Μη λειτουργικές παράμετροι όπως η ασφάλεια και η ακεραιότητα των δεδομένων.

Βάσει των ανωτέρω αποτελεί μελλοντικό θέμα ανάλυσης και μελέτης η υλοποίηση ενός μηχανισμού που δέχεται αυτές τις παραμέτρους και καθορίζει το βέλτιστο τρόπο εκτέλεσης των εργασιών (το οποίο αναφέρεται ως «*To Grid or Not To Grid*»).

8.2.3 Ο Εικονικός Υπολογιστής ως Υπηρεσία

Το συγκεκριμένο θέμα αναφέρεται στην παροχή ως υπηρεσία ενός εικονικού υπολογιστή. Η λογική του συγκεκριμένου θέματος επικεντρώνεται στην παροχή όλων των λειτουργιών ενός προσωπικού υπολογιστή ως υπηρεσίες παρεχόμενες από έναν φορέα παροχής υπηρεσιών, το οποίο αναμένεται να είναι προς όφελος των χρηστών καθώς στο συγκεκριμένο πλαίσιο ο εικονικός υπολογιστής αναφέρεται σε ένα απέραντο αριθμό υπηρεσιών και εφαρμογών. Κάθε χρήστης θα μπορεί να ζητήσει μια εφαρμογή από έναν πόρο προκειμένου να εκπληρωθούν οι υπολογιστικές / αποθηκευτικές ανάγκες του ενώ επιπλέον δεν θα είναι αρμόδιος για την επιλογή και την εγκατάσταση λογισμικού ούτε ακόμη και για την αναβάθμιση ή τη σχέση με ζητήματα ασφάλειας και συντήρησης.

Η προαναφερθείσα λειτουργικότητα απαιτεί το σχεδιασμό και την υλοποίηση δυο μερών: του εικονικού υπολογιστή που αντιλαμβάνονται οι χρήστες και του συνόλου της υποδομής που απαιτείται από την πλευρά των φορέων παροχής υπηρεσιών. Καθώς η πλευρά των παρόχων μπορεί να θεωρηθεί ήδη υλοποιημένη (παράδειγμα τέτοιας χρήσης αποτελούν οι ιστοσελίδες), θα πρέπει να σχεδιαστεί ένα λειτουργικό σύστημα από την πλευρά των χρηστών που να επιτρέπει την «απομακρυσμένη εκτέλεση» των εφαρμογών. Το γενικότερο παραπάνω σχήμα προσφέρει πολλά πλεονεκτήματα όπως για παράδειγμα τις χρεώσεις (οι χρήστες θα χρεώνονται για τη χρήση ενός λογισμικού και όχι για την αγορά του), τις αναβαθμίσεις (θα γίνονται στην πλευρά των παρόχων και με διαφανή τρόπο για τους χρήστες), τα ζητήματα συντήρησης κ.α.

Γλωσσάριο

Στον παρακάτω πίνακα παρατίθενται οι όροι που χρησιμοποιήθηκαν στη Διατριβή:

API	Application Programming Interface
ASP	Application Service Providers
B2B	Business to Business
BPEL	Business Process Execution Language
CPU	Central Processing Unit
EGA	Enterprise Grid Alliance
FIFO	First-In-First-Out
GARA	Globus Architecture for Reservation and Allocation
GGF	Global Grid Forum
GRACE	Grid Architecture for Computational Economy
GTv4	Globus Toolkit version 4
HTTP	HyperText Transfer Protocol
HTML	HyperText Markup Language
ICT	Information and Communication Technologies
IEEE	Institute of Electrical and Electronics Engineers
IP	Internet Protocol
ISO	International Standards Organization
IT	Information Technology
LAN	Local Area Network
MAC	Media Access Control
MFLO	Million Floating Operations
MFLOPS	Million Floating Operations Per Second
MOWS	Management of Web Services
NextGRID	Next Generation Grid Architecture Integrated Project
NGG	Next Generation Grid
OASIS	Organization for the Advancement of Structured Information Standards
OGF	Open Grid Forum
OGSA	Open Grid Services Architecture
OWL	Web Ontology Language
PC	Personal Computer
P2P	Peer to Peer
RDF	Resource Description Framework

RDQL	Query RDF Language
RPC	Remote Procedure Call
QoS	Quality of Service
SDK	Software Development Kits
SLA	Service Level Agreements
SOA	Service Oriented Architecture
SOAP	Simple Object Access Protocol
SOKU	Service Oriented Knowledge Utilities
SP	Service Provider
VIM	Virtual Infrastructure Model
VO	Virtual Organization
VPN	Virtual Private Network
URI	Universal Resource Identifier
WfMC	Workflow Management Coalition
WS	Web Service
WSDL	Web Services Description Language
WSRF	Web Services Resource Framework
XML	Extensible Markup Language

Βιβλιογραφικές Αναφορές

- [1] IBM, Grid Computing: Past, Present and Future, “An Innovation Perspective”, June, 2006, <http://www-03.ibm.com/grid/pdf/innovperspective.pdf>
- [2] I. Foster and C. Kesselman, “The Grid: Blueprint for a Future Computing Infrastructure”, Morgan Kaufmann Publishers, USA, 1999.
- [3] I. Foster, C. Kesselman, S. Tuecke, “The Anatomy of the Grid: Enabling Scalable Virtual Organizations” International Journal Supercomputer Applications, Vol. 15, No. 3, 2001.
- [4] 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.
- [5] A. Mayer, S. McGough, N. Furmento, W. Lee, M. Gulamali, S. Newhouse, and J. Darlington, “Workflow Expression: Comparison of Spatial and Temporal Approaches”, Workflow in Grid Systems Workshop, GGF-10, Berlin, March 9, 2004.
- [6] D.P. Spooner, J. Cao, S. A. Jarvis, L. He, and G. R. Nudd, “Performance-aware Workflow Management for Grid Computing”, The Computer Journal, Oxford University Press, London, UK, 2004.
- [7] Workflow Management Coalition, Terminology & Glossary, Document Number WFMC-TC-1011, Issues 3.0, February 1999.
- [8] I. Altintas, A. Birnbaum, K. Baldridge, W. Sudholt, M. Miller, C. Amoreira, Y. Potier, and B. Ludaescher, “A Framework for the Design and Reuse of Grid Workflows”,

- International Workshop on Scientific Applications on Grid Computing (SAG'04), LNCS 3458, Springer, 2005.
- [9] E. Deelman, J. Blythe, Y. Gil, and C. Kesselman, “Workflow Management in GriPhyN”, The Grid Resource Management, Kluwer, Netherlands, 2003.
- [10] E. Deelman, J. Blythe, Y. Gil, C. Kesselman, G. Mehta, S. Patil, M. H. Su, K. Vahi, M. Livny, “Pegasus: Mapping Scientific Workflow onto the Grid”, Across Grids Conference 2004, Nicosia, Cyprus, 2004.
- [11] B. Ludäscher, I. Altintas, and A. Gupta, “Compiling Abstract Scientific Workflows into Web Service Workflows”, 15th International Conference on Scientific and Statistical Database Management, Cambridge, Massachusetts, USA., IEEE CS Press, pp. 241-244, Los Alamitos, CA, USA., July 09-11, 2003
- [12] Jia Yu, Rajkumar Buyya, “A Taxonomy of Workflow Management Systems for Grid Computing”, Journal of Grid Computing, Volume 3, Issue 3 - 4, pp. 171-200, Springer, 2005
- [13] 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, pp. 98-105. IEEE Press, 2005
- [14] W. van der Aalst, A. ter Hofstede, B. Kiepuszewski, and A. Barros, “Workflow patterns”. Distributed and Parallel Databases, 14(1):551, 2003
- [15] <http://www.w3.org/Submission/OWL-S/>
- [16] Saez S., Vila J., Crespo A., A dynamic real-time scheduler for shared memory multiprocessors. Proceedings of the Eighth Euromicro Workshop on Real-Time Systems, 1996. Page(s): 158 –163.
- [17] GRIA, Grid Resources for Industrial Applications, www.gria.org

- [18] G. Bochmann and A. Hafid. “Some Principles for Quality of Service Management”, Technical report, Universite de Montreal, 1996.
- [19] R. J. Al-Ali, K. Amin, G. von Laszewski, O. F. Rana, D. W. Walker, M. Hategan, N. J. Zaluzec: “Analysis and Provision of QoS for Distributed Grid Applications”, pp. 163-182, Journal of Grid Computing, 2004
- [20] Padgett, J., K. Djemame, and P. Dew, “Grid-based SLA Management”, Lecture Notes in Computer Science, pp. 1282-1291, 2005
- [21] I. Foster, C. Kesselman, C. Lee, B Lindell, K. Nahrstedt, A. Roy, “A Distributed Resource Management Architecture that Supports Advance Reservation and Co-Allocation”, Proceedings of the International Workshop on QoS, pp.27-36, 1999
- [22] Min-You Wu, Wei Shu, H. Zhang, “Segmented min-min: a static mapping algorithm for meta-tasks on heterogeneous computing systems”, Heterogeneous Computing Workshop, 2000
- [23] Shanshan Song, Yu-Kwong Kwok, Kai Hwang, “Security- Driven Heuristics and A Fast Genetic Algorithm for Trusted Grid Job Scheduling Parallel and Distributed Processing Symposium”, 19th IEEE International 04-08 April 2005
- [24] R. Buyya, M. Murshed, D. Abramson, “A Deadline and Budget Constrained Cost-Time Optimization Algorithm for Scheduling Task Farming Applications on Global Grids”, Proceedings of the 2002 International Conference on Parallel and Distributed Processing Techniques and Applications(PDPTA702), 2002
- [25] R. Buyya, D. Abramson, S. Venugopal, “The Grid Economy”, Proceedings of the IEEE Volume 93, Issue 3, pp. 698-714, Mar 2005
- [26] Li Kenli, Tang Xiaoyong, Zhaohuan, “Grid Classified Optimization Scheduling Algorithm under the Limitation of Cost and Time”, Proceedings of the Second

International Conference on Embedded Software and Systems (ICESS'05), 0-7695-2512-1/05 IEEE

- [27] Jorge Cardoso, Amit Sheth and John Miller, "Workflow Quality of Service", Proceedings of the International Conference on Enterprise Integration and Modeling Technology and International Enterprise Modeling Conference (ICEIMT/IEMC'02), Kluwer Publishers, April 2002
- [28] J. Cardoso, J. Miller, A. Sheth, J. Arnold, "Modeling Quality of Service for Workflows and Web Service Processes", Technical Report, LSDIS Lab, Department of Computer Science University of Georgia, 2002
- [29] D. Angulo, I. Foster, C. Liu, and L. Yang, "Design and Evaluation of a Resource Selection Framework for Grid Applications", Proceedings of IEEE International Symposium on High Performance Distributed Computing (HPDC-11), Edinburgh, Scotland, 2002
- [30] R. Buyya, D. Abramson, J. Giddy, "An Economy Driven Resource Management Architecture for Computational Power Grids", International Conference on Parallel and Distributed Processing Techniques and Applications, Las Vegas, USA., 2000
- [31] R. Buyya, D. Abramson, J. Giddy, "High Nimrod-G: an architecture for a resource management and scheduling system in a global computational grid", Performance Computing in the Asia- Pacific Region, Proceedings of the Fourth International Conference Exhibition on Volume 1, pp. 283-289, 14-17 May 2000
- [32] DAGMan, <http://www.cs.wisc.edu/condor/dagman/>
- [33] C. Weng, X. Lu, "Heuristic scheduling for bag-of-tasks applications in combination with QoS in computational grid", Future Generation Computer Systems, pp. 271-280, 2005

- [34] V. Subramani, R. Ketimuthu, S. Srinivasan, P. Sadayappan, “Distributed job scheduling on computational grids using multiple simultaneous requests”, Proceedings of the 11th IEEE International Symposium on High Performance Distributed Computing, pp. 359-367, Edinburgh, Scotland, 2002
- [35] Open Grid Services Architecture (OGSA), www.ggf.org/documents/GFD.30.pdf
- [36] WSRF, The Web Services Resource Framework (WSRF) v1.2, <http://www.oasis-open.org/committees/download.php/17833/wsrf-1.2-os.zip>
- [37] Air Renderer: <http://www.sitexgraphics.com/html/air.html>
- [38] RenderMan Interface Specification v3.2, https://renderman.pixar.com/products/rispec/rispec_pdf/RISpec3_2.pdf
- [39] K. Dolkas, D. Kyriazis, A. Menychtas, T. Varvarigou, “e-Business Applications on the Grid: a toolkit for centralized workload prediction and access”, Concurrency and Computation: Practice and Experience, Wiley Interscience, 2006
- [40] C. Skianis, G. Xilouris, G. Kormentzas, A. Kourtis, “A Testbed Environment for Validation of End-to-End QoS Provision for the Content Delivery Chain over Heterogeneous Systems”, Proceedings of the 2nd International Working Conference on Performance Modelling and Evaluation of Heterogeneous Networks (HET-NETs’04), Ilkley, pp. 89/1-89/8, 2004
- [41] Pawel Plaszczak and Richard Wellner, Jr., Grid Computing – The Savvy’s Manager Guide, Morgan Kaufmann Publishers, 2006, ISBN-13:978-0-1274-2503-0
- [42] Dimosthenis Kyriazis, Konstantinos Dolkas, Andreas Menychtas and Theodora Varvarigou, “A new Workflow Mapping Mechanism for Grids”, IST Mobile & Wireless Communication Summit 2006, Myconos - Greece, 04-08 June 2006.
- [43] Dimosthenis Kyriazis, Konstantinos Tserpes, Andreas Menychtas, Antonios Litke and Theodora Varvarigou, “An innovative Workflow Mapping Mechanism for Grids in

- the frame of Quality of Service”, Future Generation Computer Systems (The International Journal of Grid Computing), Elsevier, 2007.
- [44] Fran Berman, Anthony J. G. Hey, Geoffrey C. Fox, Grid Computing – Making the Global Infrastructure A Reality, John Wiley and Sons Ltd, 2003, ISBN 0-470-85319-0
- [45] A. Sahai and S. Graupner and V. Machiraju and A. Moorsel. (Specifying and Monitoring) Guarantees in Commercial Grids through SLA. Proceedings of the 3rd IEEE/ACM CCGrid2003, 2003.
- [46] K. Czajkowski and I. Foster and C. Kesselman and V. Sander and S. Tuecke SNAP: A Protocol for Negotiating Service Level Agreements and Coordinating Resource Management in Distributed Systems. Proceedings of the 8th Workshop on Job Scheduling Strategies for Parallel Processing, 2002.
- [47] Web Services Description Language (WSDL) Version 2.0, <http://www.w3.org/TR/wsdl20/>
- [48] The Business Grid: Providing Transactional Business Processes via Grid Services Frank Leymannand, Kai Güntzel.
- [49] Open Grid Forum, www.ogf.org
- [50] The World Wide Web Consortium, www.w3.org
- [51] Grimoires model, <http://twiki.grimoires.org/bin/view/Grimoires/>
- [52] RDQL - A Query Language for RDF, <http://www.w3.org/Submission/RDQL/>
- [53] Peer Hasselmeyer, Changtao Qu, Lutz Schubert, Bastian Koller, and Philipp Wieder. Towards Autonomous Brokered SLA Negotiation. In: Proceedings of the eChallenges e-2006 Conference, Barcelona, Spain, October 2006.
- [54] Konstantinos Tserpes, Dimosthenis Kyriazis, Andreas Menychtas, Theodora Varvarigou, Fabrizio Silvestri, Domenico Laforenza, “Business Environments and QoS:

- Correlating Customer Requirements with Provider Capabilities”, CoreGRID’07 Symposium, 2007.
- [55] Peer Hasselmeyer, Bastian Koller, Lutz Schubert, Philipp Wieder, “Towards SLA - Supported Resource Management”, HPCC Conference Proceedings, 2006.
- [56] Duan-Shin Lee, Chun-Min Chen, Weighted fair queueing and compensation techniques for wireless packet switched networks. The 5th International Symposium on Wireless Personal Multimedia Communications, 2002. Volume: 3, 27-30 Oct. 2002 Page(s): 887 -891.
- [57] Bennett, J.C.R., Hui Zhang, WF2Q: worst-case fair weighted fair queueing INFOCOM '96. Proceedings IEEE on Fifteenth Annual Joint Conference of the IEEE Computer Societies. Networking the Next Generation. Volume: 1, 24-28 March 1996 Page(s): 120 -128 vol.1.
- [58] Peha, J.M., Tobagi, F.A., Evaluating scheduling algorithms for traffic with heterogeneous performance objectives. Global Telecommunications Conference, 1990, and Exhibition. 'Communications: Connecting the Future', GLOBECOM '90, IEEE , 2-5 Dec. 1990 Page(s): 21 -27 vol.1
- [59] Andreas Menychtas, Konstantinos Dolkas, Dimosthenis Kyriazis and Theodora Varvarigou, “Building Portals to access Grid Middleware”, GGF 14 Conference Chicago, USA, 27-30 June 2005.
- [60] Spuri M, Buttazzo G.C., Efficient aperiodic service under earliest deadline scheduling. Real-Time Systems Symposium, 1994, Proceedings. , 7-9 Dec. 1994 Page(s): 2 –11.
- [61] Sung-Heun Oh, Seung-Min Yang, A Modified Least-Laxity-First scheduling algorithm for real-time tasks. Proceedings. Fifth International Conference on Real-Time Computing Systems and Applications, 1998., 27-29 Oct. 1998 Page(s): 31 –36.

- [62] Kreuzinger J., Schulz A., Pfeffer M., Ungerer T., Brinkschulte U., Krakowski C., “Real-time scheduling on multithreaded processors”. Seventh International Conference on Real-Time Computing Systems and Applications, 2000., 12-14 Dec. 2000 Page(s): 155 –159.
- [63] Dimosthenis Kyriazis, Konstantinos Tserpes, Andreas Menychtas, Ioannis Sarantidis and Theodora Varvarigou, “Service Selection and Workflow Mapping for Grids: An approach exploiting Quality of Service Information”, Concurrency and Computation: Practice and Experience, Willey Interscience, 2007.
- [64] Al-Mouhamed M.A., Lower bound on the number of processors and time for scheduling precedence graphs with communication costs. IEEE Transactions on Software Engineering, Volume: 16 Issue: 12, Dec. 1990 Page(s): 1390 –1401.
- [65] NextGRID Integrated Project, <http://www.nextgrid.org>
- [66] Konstantinos Tserpes, Dimosthenis Kyriazis, Andreas Menychtas and Theodora Varvarigou, “A Novel Mechanism for Provisioning of High-Level Quality of Service Information in Grid Environments”, Special Issue on “Performance Evaluation of QoS-aware Heterogeneous Systems, European Journal of Operational Research, 2007.
- [67] Haddad E., Optimal load sharing in dynamically heterogeneous systems. Proceedings. Seventh IEEE Symposium on Parallel and Distributed Processing, 1995. , 25-28 Oct. 1995 Page(s): 346 –353.
- [68] Antonios Litke, Dimitrios Halkos, Konstantinos Tserpes, Dimosthenis Kyriazis and Theodora Varvarigou, “Fault Tolerant and Prioritized Scheduling in OGSA-based Mobile Grids”, Concurrency and Computation: Practice and Experience, Willey Interscience, 2007.

- [69] Tei-Wei Ku, Wang-Ru Yang, Kwei-Jay Lin, A class of rate-based real-time scheduling algorithms. IEEE Transactions on Computers, Volume: 51 Issue: 6, June 2002
Page(s): 708 –720.
- [70] Jeffay K., Stanat D.F., Martel C.U., On non-preemptive scheduling of period and sporadic tasks. Real-Time Systems Symposium, 1991. Proceedings. Twelfth, 4-6 Dec. 1991
Page(s): 129-139.
- [71] A. Menychtas, D. Apostolopoulos, D. Kyriazis, K. Christodoulopoulos, H. Avramopoulos and Theodora Varvarigou, “Enabling a Network Simulation Application on Grid Infrastructure”, 11th Panhellenic Conference on Informatics (PCI 2007), Patras, Greece, 2007.
- [72] Next Generation Grid Experts Group Report, <http://cordis.europa.eu/ist/grids/ngg.htm>
- [73] Altintas, C. Berkley, E. Jaeger, M. Jones, B. Ludscher, and S. Mock. “Kepler: Towards a Grid-Enabled system for scientific workflows”. In Workflow in Grid Systems Workshop in GGF10, Berlin, March 2004
- [74] Y. Gao, H. Rong, J. Z. Huang, “Adaptive grid job scheduling with genetic algorithms”, Future Generation Computer Systems 21, pp. 151–161, 2005
- [75] Foster, C. Kesselman, C. Lee, B Lindell, K. Nahrstedt, A. Roy , “A Distributed Resource Management Architecture that Supports Advance Reservation and Co-Allocation”, Proceedings of the International Workshop on QoS, pp.27-36, 1999
- [76] Al-Ali, R.; Rana, O.; Walker, D.; Jha, S. & Sohail, S. “G-QoS: Grid Service Discovery Using QoS Properties.” Computing and Informatics Journal, 21 (4), 2002. 363-82.
- [77] Open Grid Forum, <http://www.ogf.org>
- [78] Enterprise Grid Alliance (EGA) <http://www.gridalliance.org/>
- [79] Global Grid Forum, <http://www.ggf.org>

- [80] K. Ramamritham, J.A.Stankovic, and P.-F. Shiah, "Efficient Scheduling Algorithms for Real-time Multiprocessor Systems," IEEE Trans. on Parallel and Distributed Systems, vol.1, no.2, pp.184-194, Apr. 1990.
- [81] F. Wang, K. Ramamritham, J.A. Stankovic, "Determining redundancy levels for fault tolerant real-time systems", IEEE Trans. Computers, vol 44, 1995 Feb.
- [82] Scheduling Working Group of the Grid Forum, Document: 10.5, September 2001.
- [83] D. Abramson, R. Sasic, J.Giddy and B. Hall, "Nimrod: A Tool for Performing Parametised Simulations using Distributed Workstations", The 4th IEEE Symposium on High Performance Distributed Computing, Virginia, August 1995.
- [84] The Unicore project, <http://www.unicore.org/forum.htm>
- [85] F. Berman, A. Chien, K. Cooper, J. Dongarra, I. Foster, D. Gannon, L. Johnsson, K. Kennedy, C. Kesselman, J. Mellor-Crummey, D.Reed, L. Torczon, and R. Wolski. "The GrADS Project: Software Support for High-Level Grid Application Development". International Journal of High Performance Computing Applications, Winter 2001 (Volume 15, Number 4), pp. 327-344.
- [86] J. Frey, T. Tannenbaum, I. Foster, M. Livny, and S. Tuecke, "Condor-G: A computation management agent for multiinstitutional grids", Cluster Computing 5 (2002), 237-246.
- [87] M. Faerman, S. Figueira, J. Hayes, G. Obertelli, J. Schopf, G. Shao, S. Smallen, N. Spring, A. Su, D. Zagorodnov, "Adaptive Computing on the Grid Using AppLeS", IEEE Transactions on Parallel and Distributed Systems, Volume 14 , Issue 4 (April 2003), p. 369 – 382
- [88] D. Nurmi, J. Brevik, R. Wolski "Modeling Machine Availability in Enterprise and Wide-area Distributed Computing Environments" UCSB Computer Science Technical Report Number CS2003-28.

- [89] Xiaoshan He, Xian-He Sun, and Gregor von Laszewski, "QoS Guided Min-Min Heuristic for Grid Task Scheduling", *Journal of Computer Science and Technology*, Special Issue on Grid Computing, 18(4), 2003.
- [90] L. Carrington, A. Snavely, N. Wolter, "A performance prediction framework for scientific applications", *Future Generation Computer Systems* 22 (2006) 336–346.
- [91] R. Wolski, N. Spring, and J. Hayes, "The Network Weather Service: A Distributed Resource Performance Forecasting Service for Metacomputing", *Future Generation Computer Systems*, Vol. 15, (1999), pp. 757-768.
- [92] Jahanzeb Sherwani, Nosheen Ali, Nausheen Lotia, Zahra Hayat, and Rajkumar Buyya, "Libra: A Computational Economy based Job Scheduling System for Clusters", *International Journal of Software: Practice and Experience*, Volume 34, Issue 6, Pages: 573-590, Wiley Press, USA, May 2004.
- [93] Dave Snelling, Malcolm Atkinson, Sven van den Berghe, Konstantinos Dolkas, Mark Gilbert, Alastair Hume, Greg Kohring, Guy Lonsdale, Kostas Tserpes, Thomas Sandholm, Mike Surridge, *NextGRID Architecture*, January 17, 2005, http://www.nextgrid.org/download/publications/NextGRID_Conceptual_Architecture_V_1_5.pdf
- [94] David Snelling, Mike Fisher, Achim Basermann, *NextGRID Vision and Architecture White Paper*, 2005
- [95] D. Snelling, A. Anjomshoaa, "NextGRID Architectural Concepts", *CoreGRID'07 Symposium*, 2007.
- [96] Foster, I., Kishimoto, H., Savva, A., Berry, D., Djaoui, A., Grimshaw, A., Horn, B., Maciel, F., Siebenlist, F., Subramaniam, R., Treadwell, J., and Von Reich, J.: *The Open Grid Services Architecture*, Version 1.0. Global Grid Forum OGSA-WG. GFD-I.030, January 2005, <http://www.ggf.org/documents/GWD-I-E/GFD-I.030.pdf>.

- [97] K-Wf grid Knowledge-based Workflow System for Grid Applications,
<http://www.kwfgrid.net>
- [98] Ewa Deelman, James Blythe, Yolanda Gil, and Carl Kesselman, Grid Resource Management, chapter Workflow Management in GriPhyN. Kluwer, 2003. Available from
http://www.isi.edu/~deelman/Pegasus/grm_chapter.pdf
- [99] Stefano Beco, Barbara Cantalupo, Ludovico Giammarino, Mike Surridge and Nikolaos Matskanis. OWL-WS: A Workflow Ontology for Dynamic Grid Service Composition, accepted to the 1st IEEE International Conference on e-Science and Grid Computing, Dec. 5 - 8, 2005, Melbourne, Australia
- [100] R. Buyya, J. Giddy, D. Abramson, An Evaluation of Economy-based Resource Trading and Scheduling on Computational Power Grids for Parameter Sweep Applications, The Second Workshop on Active Middleware Services, In conjunction with HPDC 2001, 2000, Pittsburgh, USA (Kluwer Academic Press)
- [101] R. Buyya, D. Abramson, and J. Giddy, A Case for Economy Grid Architecture for Service-Oriented Grid Computing, 10th IEEE International Heterogeneous Computing Workshop, with IPDPS 2001, SF, California, USA, April 2001.
- [102] Stefano Beco, Barbara Cantalupo, Annalisa Terracina, “The Role of Workflow in Next Generation Business Oriented Grids: Two Different Approaches Leading to a Unified Vision”, eScience Conference Proceedings, 2006