



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

**Διαχείριση Περιεχομένου με Χρήση Κατανεμημένου
Συστήματος Αρχείων σε Πληροφοριοκεντρικά
Δίκτυα**

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Άγγελος-Χρήστος Γ. Αναδιώτης

Αθήνα, Ιανουάριος 2014



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

**Διαχείριση Περιεχομένου με Χρήση Κατανεμημένου
Συστήματος Αρχείων σε Πληροφοριοκεντρικά
Δίκτυα**

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Άγγελος-Χρήστος Γ. Αναδιώτης

Συμβουλευτική Επιτροπή: **Ιάκωβος Στ. Βενιέρης**
Δήμητρα-Θεοδώρα Ι. Κακλαμάνη
Νικόλαος Κ. Ουζούνογλου

Εγκρίθηκε από την επταμελή εξεταστική επιτροπή την 27η Ιανουαρίου 2014

.....
Ι. Στ. Βενιέρης
Καθηγητής Ε.Μ.Π.

.....
Δ.-Θ. Ι. Κακλαμάνη
Καθηγήτρια Ε.Μ.Π.

.....
Ν. Κ. Ουζούνογλου
Καθηγητής Ε.Μ.Π.

.....
Σ. Παπαβασιλείου
Αναπληρωτής
Καθηγητής Ε.Μ.Π.

.....
Χ. Ζ. Πατρικάκης
Επίκουρος Καθηγητής
Τ.Ε.Ι. Πειραιά

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

.....
Χ. Δουληγέρης
Καθηγητής Παν. Πειραιώς

Αθήνα, Ιανουάριος 2014

.....
Άγγελος-Χρήστος Γ. Αναδιώτης
Διδάκτωρ Μηχανικός Ε.Μ.Π.

Copyright © Άγγελος-Χρήστος Γ. Αναδιώτης, 2014.

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 στην κερδοσκοπική χρήση της εργασίας πρέπει να απευθύνονται προς το συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν το συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου. Ειδικότερα, η έγκριση της Διδακτορικής Διατριβής από την Ανώτατη Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του Ε. Μ. Πολυτεχνείου δεν υποδηλώνει αποδοχή των γνώμων του συγγραφέα (Ν. 5343/1932, Άρθρο 202).

*Στον πατέρα μου Γιώργο, τη μητέρα μου Δήμητρα και την αδερφή μου Ελένη.
Για τη στήριξή τους όλα αυτά τα χρόνια.*

Περίληψη

Η διαρκής αύξηση των χρηστών του διαδικτύου έχει δημιουργήσει νέες προκλήσεις σε ό,τι αφορά τη διαχείριση του διαθέσιμου περιεχομένου, το οποίο μπορεί να βρίσκεται αποθηκευμένο σε οποιοδήποτε σημείο του κόσμου. Η παρούσα διατριβή έχει ως στόχο να παράσχει μια ολοκληρωμένη λύση για τη διαχείριση του κατανεμημένου, πλέον, περιεχομένου αναγνωρίζοντας τρία επίπεδα: την οργάνωση, την ανακάλυψη και την παράδοσή του στο χρήστη. Για καθένα από αυτά τα επίπεδα, δίνεται ένα περίγραμμα των υπαρχουσών τεχνολογιών, ενώ παρουσιάζονται και νέες μέθοδοι, οι οποίες ενισχύουν τις τρέχουσες λύσεις που περιλαμβάνονται στη βιβλιογραφία, προκειμένου να αυξηθεί η απόδοση του συστήματος. Η ενοποίηση των τριών επιπέδων γίνεται με χρήση πρότυπων τεχνολογιών, η ανάπτυξη πολλών εκ των οποίων αποτέλεσε αντικείμενο έρευνας της διατριβής.

Στο επίπεδο της παράδοσης του περιεχομένου στον τελικό χρήστη, μελετάται η χρήση πληροφοριοκεντρικών δικτύων, τα οποία έχουν αποτελέσει έναν από τους βασικούς άξονες έρευνας για τη μορφή που πρόκειται να έχει στο μέλλον το διαδίκτυο. Λαμβάνοντας υπόψη ότι σήμερα, αλλά και τα προσεχή χρόνια, ο μεγαλύτερος όγκος δεδομένων διακινείται από εφαρμογές βίντεο και από διαμοιρασμό περιεχομένου σε δίκτυα ομότιμων κόμβων, γίνεται μελέτη της επίδοσης διαδεδομένων πρωτοκόλλων σε αυτούς τους τομείς, ενώ προτείνονται και τρόποι βελτίωσης της απόδοσης του δικτύου, με αντίκτυπο τόσο στους παρόχους υπηρεσιών διαδικτύου όσο και στους τελικούς χρήστες.

Στο επίπεδο της ανακάλυψης του περιεχομένου προτείνεται η χρήση πρωτοκόλλων δημοσίευσης/συνδρομής βασιζόμενων στο περιεχόμενο, το οποίο μπορεί, αφενός, να ικανοποιούν καλύτερα τις ανάγκες των χρηστών σε εκφραστικότητα των ερωτημάτων τους, αλλά, αφετέρου, αντιμετωπίζουν προβλήματα κλιμάκωσης. Η προτεινόμενη προσέγγιση επιφέρει βελτίωση της απόδοσης του συστήματος σε σχέση με παραδοσιακές λύσεις τόσο σε όρους κλιμάκωσης όσο και εξισορρόπησης φόρτου, μέσω της χρήσης μηχανισμού κατανεμημένης κρυφής μνήμης.

Στο επίπεδο της οργάνωσης του περιεχομένου παρουσιάζεται ένα κατανεμημένο σύστημα αρχείων, το οποίο θεωρεί δυναμική δομή καταλόγων, που προκύπτει από οντολογία καθορισμένη από τον ίδιο το χρήστη. Βασιζόμενο σε πρότυπες τεχνολογίες, προτείνει έναν ενιαίο τρόπο αναπαράστασης του κατανεμημένου περιεχομένου, χρησιμοποιώντας τα μεταδεδομένα της περιγραφής του που απαιτούνται για την ανακάλυψή του, ενώ η προ-

στασία του γίνεται με άδειες χρήσης, οι οποίες ελέγχονται από καταναμημένες υπηρεσίες, αποσυνδέοντας με αυτόν τον τρόπο τον πάροχο του περιεχομένου από διαδικασίες ελέγχου πρόσβασης.

Η ενοποίηση του συστήματος γίνεται με χρήση των πρότυπων δομών δεδομένων του MPEG-21 και των πρότυπων τεχνολογιών του μερισμικού MPEG-M. Τόσο το προτεινόμενο πρωτόκολλο ανακάλυψης όσο και τα πρωτόκολλα παράδοσης του περιεχομένου δύνανται να υποστηριχθούν από τα εν λόγω πρότυπα, τα οποία με τη σειρά του χρησιμοποιεί το προτεινόμενο καταναμημένο σύστημα αρχείων. Έτσι, η διατριβή επιχειρεί να δώσει μια διαφανή λύση, η οποία δεν επηρεάζει σε τίποτα τον πιο συνηθισμένο τρόπο πρόσβασης του χρήστη στην πληροφορία, που είναι μέσω ενός αρχείου.

Λέξεις κλειδιά: Πληροφοριοκεντρικά δίκτυα, πρωτόκολλο δημοσίευσης/συνδρομής, καταναμημένο σύστημα αρχείων, πρότυπες τεχνολογίες μερισμικού.

Abstract

The continuous growth of Internet users has created new challenges in managing the distributed content, which may be located anywhere around the globe. This thesis aims in providing an integrated solution for distributed content management and in this course, it identifies three layers: organization, discovery and delivery of content to the end user. For each one of these layers, an outline of the existing technologies is provided, while new methods are also identified, which improve the behavior of the current solutions in the literature and thus, the overall system performance. Integration of all the three layers is achieved by using standard technologies, to which this thesis has contributed.

In the content delivery layer, information-centric networks are studied, which have emerged as a promising technology that can be used in the Internet of the future. Considering that, both today and in the near future, video streaming and peer to peer applications are expected to consume the greater fraction of Internet bandwidth, the performance of existing protocols is evaluated, while enhancements are suggested, which can improve the performance of the information-centric network, yielding advantages for all the Internet Service Providers (ISPs) and the end users.

In the content discovery layer, the use of content-based publish/subscribe protocols is suggested, as they can better capture the users' requirements in terms of query expressiveness. On the other hand though, these protocols often have scalability issues, which are generally addressed in the literature by considering special characteristics of the underlying systems. The proposed approach improves the performance of standard algorithms employed in this context, both in terms of scalability and load balancing, by introducing the use of a distributed caching mechanism.

In the content organization layer, a novel distributed file system called EuterpeFS is presented, which considers a dynamic directory structure based on ontologies specified by the users. The use of a standard data structure is suggested to be used as a unified container of distributed content, relying only on the content's descriptive metadata required for its discovery, while access control to it is performed in a distributed manner, by means of licenses expressed with a standard language.

These three layers are integrated using MPEG-21 standard data structures and MPEG-M

standard technologies. All the suggested content discovery and the content delivery protocols can be supported by these standards, which are anyway used by EuterpeFS. In this course, this thesis provides a transparent solution in content management, without changing anything in the most common way that users access information; that is, through files.

Keywords: Information-centric networks, publish/subscribe protocol, distributed file system, standard middleware technologies.

Ευχαριστίες

Η παρούσα διατριβή αποτελεί το επιστέγασμα της ερευνητικής μου δραστηριότητας στο Εργαστήριο Ευφών Επικοινωνιών και Δικτύων Ευρείας Ζώνης του ΕΜΠ. Όλο αυτό το διάστημα, είχα τη χαρά και την τιμή να γνωρίσω και να συνεργαστώ με πολλούς ανθρώπους, τους οποίους θα ήθελα σε αυτό το σημείο να ευχαριστήσω για τη στήριξη, την καθοδήγηση και τη βοήθειά τους.

Αρχικά, θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή μου κύριο Ιάκωβο Βενιέρη, Καθηγητή ΕΜΠ, για την εμπιστοσύνη του στις ικανότητές μου, τις ευκαιρίες που μου έδωσε να ασχοληθώ με ενδιαφέροντα ερευνητικά θέματα, την καθοδήγηση, την υποστήριξη και την πίστη του σε μένα. Στη συνέχεια, την κυρία Δήμητρα Κακλαμάνη, Καθηγήτρια ΕΜΠ, για τη στήριξή της, τις συμβουλές της και την άψογη συνεργασία που είχαμε όλα αυτά τα χρόνια. Επίσης, θα ήθελα να ευχαριστήσω τον κύριο Χαράλαμπο Πατρικάκη, Επίκουρο Καθηγητή ΤΕΙ Πειραιά, για τη στενή συνεργασία και τη βοήθεια που μου προσέφερε. Επιπλέον, τον κύριο Χρήστο Δουληγέρη, Καθηγητή Πανεπιστημίου Πειραιώς, για την καθοδήγηση και τις υποδείξεις του αλλά και την τιμή που μου έκανε να είναι μέλος της επταμελούς επιτροπής μου. Τέλος, θερμές ευχαριστίες στον κύριο Νικόλαο Ουζούνου, Καθηγητή ΕΜΠ, και και στον κύριο Συμεών Παπαβασιλείου, Αναπληρωτή Καθηγητή ΕΜΠ, που πάντοτε ευγενικοί στήριξαν τις προσπάθειές μου καθώς και στον Καθηγητή ΕΜΠ, κύριο Ανδρέα-Γεώργιο Σταφυλοπάτη που μου έκανε την τιμή να συμπεριληφθεί στην επταμελή επιτροπή μου.

Ένα μεγάλο ευχαριστώ οφείλω στον Αζίζ Μούσας για τη συνεργασία, τη συμπαράσταση, την κατανόησή του και πάνω από όλα επειδή ήξερα ότι μπορώ να βασιστώ πάνω του. Στο Γιώργο Λιουδάκη, που ήταν πάντα εκεί όταν χρειάστηκα τις συμβουλές του για οποιονδήποτε λόγο. Στον Παναγιώτη Γκόνη, για την ευχάριστη συνεργασία και τις συμβουλές του. Στη Σοφία Καπελλάκη για τις άπειρες ώρες που αφιέρωσε να με ακούει. Στο Χρήστο Παππά για την εμπιστοσύνη που μου έδειξε συμπεριλαμβάνοντάς με στην ερευνητική του προσπάθεια. Στη Δέσποινα Μεριδου και τη Μαριέλα Παπαδοπούλου για την κατανόηση και την εμπιστοσύνη τους. Μαζί με τους προαναφερθέντες, σε όλα τα παιδιά του εργαστηρίου, παλιότερους και νεότερους, για τις αλησμόνητες στιγμές που περάσαμε μαζί. Σε όλους αυτούς που έχω τη χαρά να αποκαλώ πλέον φίλους μου.

Τέλος, μέσα από αυτές τις γραμμές θα ήθελα να εκφράσω τις εκ βαθέων ευχαρι-

στίες μου στην πολυαγαπημένη μου οικογένεια που δεν έπαψε στιγμή να με πιστεύει και να με στηρίζει, ακόμα και όταν εγώ δεν έβλεπα το λόγο. Στη Ζωή που στάθηκε δίπλα μου ακόμα και στις πιο σκοτεινές στιγμές μου χωρίς να χάνει την πίστη της σε μένα. Στους φίλους μου που δε με ξέχασαν αν και δεν ήμουν πάντα δίπλα τους. Χωρίς αυτούς, τίποτα από ό,τι ακολουθεί δε θα υπήρχε.

Πίνακας Περιεχομένων

	Σελ.
Περίληψη	viii
Abstract	x
Ευχαριστίες	xii
Πίνακας Περιεχομένων	xiii
Πίνακας Σχημάτων	xvii
Πίνακας Αλγορίθμων	xxi
1 Εισαγωγή	1
1.1 Παράδοση Περιεχομένου στον Τελικό Χρήστη	2
1.2 Ανακάλυψη Κατανεμημένου Περιεχομένου	3
1.3 Οργάνωση Περιεχομένου στον Τελικό Χρήστη	4
1.4 Διάρθρωση Διατριβής	5
2 Παράδοση Περιεχομένου με Πληροφοριοκεντρικά Δίκτυα	7
2.1 Εισαγωγή	7
2.2 Σχήματα Ονομάτων για Πληροφοριοκεντρικά Δίκτυα	8
2.2.1 Γενικά	8
2.2.2 Η προσέγγιση του TRIAD	9
2.2.3 Η προσέγγιση του DONA	10
2.3 Αρχιτεκτονικές Πληροφοριοκεντρικών Δικτύων	14

2.3.1	Μοντέλο Υπερκειμένου Δικτύου	14
2.3.2	Μοντέλο Καθαρού Μητρώου	15
2.4	Ο Προσομοιωτής JCCNxSim	18
2.4.1	Το Μεσισμικό Προσομοίωσης του PeerSim	18
2.4.2	Διάρθρωση και Υλοποίηση του JCCNxSim	20
2.5	Υποστήριξη Διαμοιρασμού Περιεχομένου σε Δίκτυα Ομότιμων Κόμβων	23
2.5.1	Το Πρωτόκολλο Διαμοιρασμού Περιεχομένου BitTorrent	24
2.5.2	Αρχιτεκτονική Ομότιμης Επικοινωνίας σε Πληροφοριοκεντρικό Δίκτυο	25
2.5.3	Πειραματική Εκτίμηση Επιδόσεων της Προτεινόμενης Αρχιτεκτονικής	28
2.6	Υποστήριξη Ροής Πολυμεσικού Περιεχομένου	31
2.6.1	Υποστήριξη Ροής Βίντεο πάνω από Πληροφοριοκεντρικό Δίκτυο	33
2.6.2	Εγγενής Υποστήριξη Προσαρμοσμένης Ροής Βίντεο στο Δίκτυο	35
2.6.3	Πειραματική Εκτίμηση Επιδόσεων Πρωτοκόλλων	37
3	Το Πρωτόκολλο Κατανεμημένης Ανακάλυψης Περιεχομένου CCPS	51
3.1	Εισαγωγή	51
3.2	Θεμελιώδεις Σχεδιαστικές Αρχές του CCPS	52
3.3	Τρέχουσες Τάσεις στην Ευρύτερη Ερευνητική Περιοχή	53
3.4	Περιγραφή Πρωτοκόλλου CCPS	55
3.4.1	Ανακάλυψη Περιεχομένου	55
3.4.2	Διαμοιρασμός Γνώσης	58
3.5	Θεωρητική Ανάλυση Απόδοσης Συστήματος	60
3.6	Υλοποίηση Πρωτοκόλλου CCPS	63
3.6.1	Γενική Προσέγγιση	63
3.6.2	Δημοσίευση Περιεχομένου	63
3.6.3	Συνδρομή σε Περιεχόμενο	65
3.7	Πειραματική Εκτίμηση Επιδόσεων CCPS	67
3.7.1	Διαμόρφωση Περιβάλλοντος Προσομοίωσης	67
3.7.2	Μηνύματα Ανακάλυψης Περιεχομένου	68
3.7.3	Μηνύματα Περιεχομένου	74

3.7.4	Επικάλυψη Μηνυμάτων	76
4	Υποστήριξη Ετερογένειας με τα Πρότυπα MPEG-21 και MPEG-M	79
4.1	Βασικές Δομές Δεδομένων του MPEG-21	79
4.1.1	Το Ψηφιακό Αντικείμενο	79
4.1.2	Η Γλώσσα Περιγραφής Δικαιωμάτων	83
4.1.3	Το Λεξικό Δικαιωμάτων	84
4.2	Οι Τεχνολογίες του MPEG-M	84
4.2.1	Αρχιτεκτονική του MPEG-M	85
4.2.2	Μηχανές Πρωτοκόλλου και Τεχνολογίας του MPEG-M	86
4.2.3	Στοιχειώδεις Υπηρεσίες του MPEG-M και Ενοποίησή τους	88
4.2.4	Υλοποίηση Υπηρεσίας Δημοσίευσης/Συνδρομής στο MPEG-M	90
5	Το Κατανεμημένο Σύστημα Αρχείων EuterpeFS	93
5.1	Εισαγωγή	93
5.2	Απαιτήσεις Κατανεμημένων Συστημάτων Αρχείων	94
5.2.1	Διαφάνεια	94
5.2.2	Παράλληλη Ενημέρωση Αρχείων	96
5.2.3	Αντίγραφα Αρχείων	96
5.2.4	Ετερογένεια	97
5.2.5	Ανοχή Βλαβών	97
5.2.6	Συνέπεια	97
5.2.7	Ασφάλεια	98
5.2.8	Αποδοτικότητα	98
5.3	Αρχιτεκτονική Κατανεμημένων Συστημάτων Αρχείων	98
5.3.1	Υπηρεσία Αρχείων	99
5.3.2	Υπηρεσία Καταλόγων	103
5.3.3	Προσάρτηση Απομακρυσμένων Καταλόγων	104
5.4	Μελέτες Περίπτωσης Κατανεμημένων Συστημάτων Αρχείων	105
5.4.1	Sun Network File System	105
5.4.2	Andrew File System	107

5.4.3	Lustre	108
5.5	Ψηφιακά Αντικείμενα για Αναπαράσταση Αρχείων	109
5.6	Έλεγχος Πρόσβασης στο EuterpeFS	111
5.7	Οργάνωση Αρχείων στο EuterpeFS	114
5.8	Δημοσίευση Αρχείων στο EuterpeFS	117
5.9	Συνδρομή σε Αρχεία στο EuterpeFS	120
6	Συμπεράσματα και Μελλοντικό Έργο	123
6.1	Σύνοψη και Συμπεράσματα	123
6.2	Μελλοντικές Ερευνητικές Κατευθύνσεις	125
	Βιβλιογραφία	127
	Δημοσιεύσεις	135

Πίνακας Σχημάτων

	Σελ.
1 Η Αρχιτεκτονική ενός Κόμβου CCNx	16
2 Υποστήριξη Ομότιμης Επικοινωνίας στο Δίκτυο	26
3 Παράδειγμα Αίτησης σε Ομότιμη Επικοινωνία	28
4 Πιθανότητα Εύρεσης Περιεχομένου στην Κρυφή Μνήμη	29
5 Κίνηση εκτός Αυτόνομου Συστήματος	30
6 Πιθανότητα Εύρεσης Περιεχομένου στην Κρυφή Μνήμη ως προς Αριθμό Αρχείων	31
7 Κίνηση εκτός Αυτόνομου Συστήματος ως προς Αριθμό Αρχείων	32
8 Δημοσίευση Βίντεο σε Πληροφοριοκεντρικό Δίκτυο	34
9 Υποστήριξη Ροής Βίντεο στο Δίκτυο	36
10 Πιθανότητα Εύρεσης Περιεχομένου στην Κρυφή Μνήμη για Μη-Προσαρμοσμένη Ροή Βίντεο	40
11 Κίνηση Έξω από το Αυτόνομο Σύστημα για Μη-Προσαρμοσμένη Ροή Βίντεο	40
12 Πιθανότητα Εύρεσης Περιεχομένου στην Κρυφή Μνήμη για Προσαρμοσμένη Ροή Βίντεο	41
13 Κίνηση Έξω από το Αυτόνομο Σύστημα για Προσαρμοσμένη Ροή Βίντεο	42
14 Σύγκριση Προσεγγίσεων ως προς την Πιθανότητα Εύρεσης Περιεχομένου στην Κρυφή Μνήμη	42
15 Επιρροή Επέκτασης στην Πιθανότητα Εύρεσης Περιεχομένου στην Κρυφή Μνήμη	43
16 Χρόνος Αναμονής μέχρι την Εκκίνηση Αναπαραγωγής στη Μη-Προσαρμοσμένη Ροή	44
17 Χρόνος Αναμονής κατά την Αναπαραγωγή στη Μη-Προσαρμοσμένη Ροή	45

18	Χρόνος Αναμονής για Εκκίνηση Αναπαραγωγής στην Προσαρμοσμένη Ροή (2Mbps)	45
19	Ποιότητα Αναπαραγωγής Βίντεο στην Προσαρμοσμένη Ροή (2Mbps)	46
20	Χρόνος Αναμονής για Εκκίνηση Αναπαραγωγής στην Προσαρμοσμένη Ροή (8Mbps)	47
21	Ποιότητα Αναπαραγωγής Βίντεο στην Προσαρμοσμένη Ροή (8Mbps)	47
22	Αναβαθμίσεις στην Ποιότητα Αναπαραγωγής Βίντεο (8Mbps)	48
23	Αναβαθμίσεις/Υποβαθμίσεις στην Ποιότητα Αναπαραγωγής Βίντεο (8Mbps)	49
24	Συνδρομή στο Υπερκείμενο Δίκτυο Δύο Επιπέδων Δημοσιευόντων Κόμβων .	56
25	Διαμοιρασμός Γνώσης Μεταξύ των Συνδρομητών	58
26	Μηνύματα Ανακάλυψης ως προς Δημοσιεύοντες Κόμβους (Προσομοιώσεις) .	69
27	Μηνύματα Ανακάλυψης ως προς Δημοσιεύοντες Κόμβους (Ανάλυση)	69
28	Μηνύματα Ανακάλυψης ως προς Ερωτήματα (Προσομοιώσεις)	70
29	Μηνύματα Ανακάλυψης ως προς Ερωτήματα (Ανάλυση)	70
30	Μηνύματα Ανακάλυψης για 200000 Ερωτήματα και Περιορισμένη Κρυφή Μνήμη	71
31	Βελτίωση λ στην Επίδοση του Συστήματος (Άπειρη Κρυφή Μνήμη)	72
32	Βελτίωση λ στην Επίδοση του Συστήματος με Περιορισμένη Κρυφή Μνήμη (Προσομοιώσεις)	73
33	Πιθανότητα Εύρεσης Περιεχομένου στην Κρυφή Μνήμη (Θεωρητική Ανάλυση)	73
34	Βελτίωση λ στην Απόδοση του Συστήματος για Zipf Κατανομή	74
35	Εξισορρόπηση Φόρτου στα Μηνύματα Περιεχομένου	75
36	Εξέλιξη Μηνυμάτων Περιεχομένου (Άπειρη Κρυφή Μνήμη)	75
37	Μηνύματα Περιεχομένου για 200.000 Ερωτήματα και Περιορισμένη Κρυφή Μνήμη	76
38	Επικαλυπτόμενα Μηνύματα με Άπειρη Κρυφή Μνήμη	77
39	Επικαλυπτόμενα Μηνύματα με Περιορισμένη Κρυφή Μνήμη	77
40	Αρχιτεκτονική του Μεσισμικού MPEG-M	86
41	Παράδειγμα Αρχείου Φορτώματος Υλοποιήσεων Μηχανών	89
42	Ενοποιημένη Υπηρεσία Δημοσίευσης Περιεχομένου	91

43	Η Αρχιτεκτονική του NFS	106
44	Μορφή Ψηφιακού Αντικειμένου για Αναπαράσταση Αρχείου στο EuterpeFS .	110
45	Έλεγχος Πρόσβασης στο EuterpeFS	113
46	Παράδειγμα Μουσικής Οντολογίας - Επίπεδο Εννοιών	116
47	Παράδειγμα Μουσικής Οντολογίας - Επίπεδο Στιγμιότυπων	117

Πίνακας Αλγορίθμων

	Σελ.
1 Εξυπηρέτηση Μηνύματος Multimedia Interest σε PIT και Content Store	37
2 Δημοσίευση Περιεχομένου από Κόμβο - PUBLISH_CONTENT	64
3 Λήψη Περιεχομένου από Συνδρομητή - RECV_CONTENT	64
4 Ενημέρωση Γνώσης - UPDATE_KNOWLEDGE	65
5 Συνδρομή σε Περιεχόμενο - SUBSCRIBE	66
6 Δημοσιεύων Κόμβος που Λαμβάνει Συνδρομή - RECV_SUBSCRIPTION	67

Κεφάλαιο 1

Εισαγωγή

Το περιεχόμενο αποτελεί σημείο αναφοράς για τα σύγχρονα καταναμημένα συστήματα. Η διαχείρισή του δε, έχει αποτελέσει αντικείμενο έρευνας σε πολλά επίπεδα: από την ανακάλυψή του, μέχρι την αποστολή και οργάνωσή του στην πλευρά του τελικού χρήστη. Μέσα σε αυτό το πλαίσιο έχει προταθεί στη βιβλιογραφία ένας μεγάλος αριθμός τεχνολογιών, οι οποίες φιλοδοξούν να ικανοποιήσουν συγκεκριμένες απαιτήσεις των τελικών χρηστών, στους οποίους απευθύνονται.

Επικεντρώνοντας στη διαχείριση του περιεχομένου, η παρούσα διατριβή παρουσιάζει τεχνολογίες που εκτείνονται και στα τρία προαναφερθέντα επίπεδα, επιχειρώντας μια κατακόρυφη προσέγγιση σε ολόκληρη την αλυσίδα οργάνωσης και παράδοσης καταναμημένου περιεχομένου. Το πρώτο, χαμηλότερο, επίπεδο αφορά την παράδοση της πληροφορίας στον τελικό χρήστη και για το σκοπό αυτό εξετάζεται η χρήση δικτύων νέας γενιάς που επικεντρώνονται στο περιεχόμενο. Μέσα σε αυτό το πλαίσιο, προτείνονται επεκτάσεις των υπαρχουσών αρχιτεκτονικών για αποδοτικότερη υποστήριξη ομότιμης επικοινωνίας και παράδοσης πολυμεσικού περιεχομένου. Στη συνέχεια, το δεύτερο επίπεδο, αφορά την ανακάλυψη περιεχομένου με χρήση πρωτοκόλλου δημοσίευσης/συνδρομής σε ένα υπερκείμενο δίκτυο. Το προτεινόμενο πρωτόκολλο επικεντρώνεται σε συστήματα βασιζόμενα στο περιεχόμενο και, ακολουθώντας μια γενική προσέγγιση, βελτιώνει τη συμπεριφορά των επιδημικών αλγορίθμων που χρησιμοποιούνται ευρέως για αυτό το σκοπό. Το τρίτο επίπεδο έχει να κάνει με την προσωποποιημένη διαχείριση του περιεχομένου στην πλευρά του τελικού χρήστη και για το σκοπό αυτό περιγράφεται ένα καταναμημένο σύστημα αρχείων, το οποίο έχει τη δυνατότητα να προσαρμόζεται δυναμικά στις προτιμήσεις του εκάστοτε χρήστη, μέσω χρήσης σημασιολογικών τεχνολογιών. Τέλος, λαμβάνοντας υπόψη το μεγάλο βαθμό ετερογένειας που παρατηρείται στα σύγχρονα καταναμημένα συστήματα, και τα τρία επίπεδα ενοποιούνται μέσω της χρήσης πρότυπων τεχνολογιών, η εξέλιξη των οποίων διαμορφώθηκε εν μέρει στα πλαίσια της παρούσας διατριβής.

1.1 Παράδοση Περιεχομένου στον Τελικό Χρήστη

Ξεκινώντας από το χαμηλότερο επίπεδο της μεταφοράς και παράδοσης της πληροφορίας στον τελικό χρήστη, τα τελευταία χρόνια έχει σημειωθεί αξιοσημείωτη πρόοδος. Τα παραδοσιακά δίκτυα, συμπεριλαμβανομένου του διαδικτύου στη σημερινή του μορφή, ακολουθούν μια λογική επικεντρωμένη στην εξεύρεση της τοποθεσίας της πληροφορίας, ώστε η τελευταία να ζητηθεί από το σωστή οντότητα. Εντούτοις, από μόνο του το πρόβλημα της διευθυνσιοδότησης, πάνω στην οποία έχουν οικοδομηθεί τα σύγχρονα δίκτυα, αδυνατεί να ικανοποιήσει την απαίτηση του τελικού χρήστη για περιεχόμενο, καθώς η απευθείας χρήση των διευθύνσεων δε θα μπορούσε να κλιμακώσει σε ένα καταναμημένο περιβάλλον, ολόένα εμπλουτιζόμενο με νέες συσκευές.

Με τη χρήση του Domain Name System (DNS) για την αντιστοίχιση ονομάτων με διευθύνσεις δόθηκε μεγάλη ώθηση στη χρήση του διαδικτύου, καθώς πλέον οι χρήστες ήταν σε θέση να ζητούν το περιεχόμενο από κάποια συσκευή του δικτύου, η οποία πλέον έχει συγκεκριμένο, μοναδικό και (εν γένει) κατανοητό όνομα. Για αυτό το λόγο, μέσα σε αυτά τα χρόνια από τη σύλληψή του, το DNS έχει εξελιχθεί σε αναπόσπαστο κομμάτι της διαδικτυακής υποδομής.

Από μόνο του όμως το DNS δε θα μπορούσε να βοηθήσει στη μεγάλη εξάπλωση του διαδικτύου και δει του παγκόσμιου ιστού (Word Wide Web - WWW), καθώς η εξεύρεση της πηγής που το διαθέτει δεν είναι αρκετή. Στην πραγματικότητα, η μεγάλη επιτυχία του παγκόσμιου ιστού, που έχει ανεβάσει τόσο πολύ τη χρήση του διαδικτύου, κρύβεται εν πολλοίς στο πρωτόκολλο HTTP, το οποίο αποτελεί μια πρότυπη, ολοκληρωμένη λύση ώστε οι τελικοί χρήστες να μπορούν να ζητούν με έναν ενιαίο τρόπο τη συγκεκριμένη πληροφορία που χρειάζονται.

Σήμερα, η χρήση του HTTP δεν περιορίζεται όμως στις εφαρμογές του παγκόσμιου ιστού. Μια από τις πλέον ανερχόμενες αρχιτεκτονικές ροής πολυμεσικού περιεχομένου, το οποίο αναμένεται να κυριαρχήσει στη χρήση του διαδικτύου, βασίζεται σε αυτό το πρωτόκολλο. Όπως θα αναλυθεί και σε επόμενη ενότητα αφιερωμένη σε αυτό το σκοπό, διάφορα μοντέλα έχουν προταθεί, με ίσως πιο γνωστά το Apple HTTP Live Streaming και το MPEG Dynamic Adaptive Streaming over HTTP (DASH). Η επιτυχία τους, που αποδεικνύεται από τη διάδοσή τους, σε σχέση με τις υπάρχουσες προσεγγίσεις, οι οποίες αγνοούν το περιεχόμενο (π.χ. Real-time Transport Protocol - RTP), αποτελεί ενδεικτικό της τάσης να δημιουργούνται συστήματα που αξιοποιούν το γεγονός ότι η διακινούμενη πληροφορία στο δίκτυο φέρει κάποιο μοναδικό όνομα, που την αναγνωρίζει μονοσήμαντα στις διάφορες εφαρμογές.

Λαμβάνοντας υπόψη τα παραπάνω, δε θα μπορούσε παρά να θεωρείται φυσικό το επιχειρούμενο επόμενο βήμα στο επίπεδο του δικτύου, δηλαδή της δρομολόγησης με βάση το περιεχόμενο αυτό καθαυτό. Τα πληροφοριοκεντρικά δίκτυα (Information-Centric Networks - ICNs) φιλοδοξούν να αποτελέσουν ένα μοντέρνο μοντέλο επικοινωνίας, η οποία

βασίζεται μονάχα στο όνομα του περιεχομένου που την αντιπροσωπεύει. Οι προκλήσεις που επισύρει αυτή η νέα προσέγγιση αφορούν εν πρώτοις τον τρόπο ονοματοδοσίας και, σε επέκταση, δρομολόγησης του περιεχομένου πάνω από ένα δίκτυο, του οποίου οι κόμβοι δε φέρουν πια μια μοναδική διεύθυνση, όπως στα παραδοσιακά δίκτυα TCP/IP.

Με μια προσεκτικότερη, όμως, ματιά διαφαίνεται και ένα ακόμα πλεονέκτημα μιας πληροφοριοκεντρικής αρχιτεκτονικής. Συγκεκριμένα, λαμβάνοντας υπόψη και την ευρεία και ολοένα αυξανόμενη χρήση των δικτύων παράδοσης περιεχομένου (Content Delivery Networks - CDNs), τα πληροφοριοκεντρικά δίκτυα είναι σε θέση να διατηρούν κρυφή μνήμη, κατανεμημένη σε όλους τους δρομολογητές που υποστηρίζουν τα πρωτόκολλά τους. Έτσι, δύναται να αποφορτιστεί το δίκτυο και μάλιστα μέσω εγγενούς υποστήριξης μηχανισμών κρυφής μνήμης, χωρίς απαραίτητα να απαιτούνται ακριβές λύσεις που παρέχουν τρίτες οντότητες, όπως γίνεται σήμερα με τα CDNs.

Η επιτυχία ενός νέου μοντέλου δικτυακής επικοινωνίας όμως, συνδέεται σε μεγάλο βαθμό με την ενσωμάτωση υπάρχουσών τεχνολογιών, οι οποίες ήδη χρησιμοποιούνται ευρέως. Και, αν για τον παγκόσμιο ιστό η μετάβαση μπορεί να είναι μάλλον τετριμμένη, αφού σε κάθε περίπτωση ο χρήστης ζητά περιεχόμενο με το όνομά του, δεν ισχύει το ίδιο για τη ροή πολυμεσικού περιεχομένου και, ακόμα περισσότερο, για την ομότιμη επικοινωνία. Σε ό,τι αφορά το πρώτο, έχει ήδη γίνει αρκετό έργο, κυρίως λαμβάνοντας υπόψη πρωτόκολλα με βάση το HTTP. Σε αυτή τη διατριβή παρουσιάζονται κάποια ζητήματα που ανακύπτουν από τη χρήση αυτών των μηχανισμών, ως προς την πλήρη εκμετάλλευση των νέων δυνατοτήτων που παρέχουν τα πληροφοριοκεντρικά δίκτυα και παρουσιάζεται ένας νέος μηχανισμός για την προσαρμοσμένη ροή, η οποία ενέχει πλεονεκτήματα τόσο για τον πάροχο υπηρεσιών διαδικτύου, όσο και για τον τελικό χρήστη.

Στον τομέα των ομότιμων δικτύων, όμως, τα πράγματα είναι διαφορετικά. Δεδομένου ότι στα πληροφοριοκεντρικά δίκτυα, το δίκτυο γνωρίζει το περιεχόμενο και κάνει την κατάλληλη δρομολόγηση στον πάροχο, όταν πρόκειται για ομότιμη επικοινωνία, όπου ο κάθε κόμβος πρέπει να ζητήσει ρητά το περιεχόμενο ρητά από κάποιον άλλο, το δίκτυο αποκτά δευτερεύοντα ρόλο, δρομολογώντας απλά προς μια υπηρεσία και χωρίς να μπορεί να αξιοποιήσει την κρυφή του μνήμη. Για να αντιμετωπιστεί αυτό το φαινόμενο, παρουσιάζεται ένας μηχανισμός, ο οποίος αφενός διατηρεί τις αρχές της ομότιμης επικοινωνίας και αφετέρου δίνει τη δυνατότητα στο δίκτυο να βελτιώσει την επίδοση του συστήματος μέσω του μηχανισμού κρυφής μνήμης του.

1.2 Ανακάλυψη Κατανεμημένου Περιεχομένου

Η μελέτη ενός κατανεμημένου συστήματος ανακάλυψης περιεχομένου μπορεί να γίνει είτε από τη σκοπιά ενός πρωτοκόλλου αναζήτησης είτε ενός πρωτοκόλλου δημοσίευσης/συνδρομής. Η συνήθης διαφοροποίηση των δύο αυτών προσεγγίσεων είναι ότι στα μεν

πρώτα, επιστρέφεται μόνο το περιεχόμενο που είναι διαθέσιμο τη στιγμή της αναζήτησης, ενώ στα δεύτερα επιστρέφεται το περιεχόμενο που θα δημοσιευτεί μελλοντικά. Στην παρούσα διατριβή γίνεται (όχι για πρώτη φορά) μια ενοποίηση των δύο αυτών προσεγγίσεων, παρέχοντας μια ολοκληρωμένη λύση για ένα σύστημα δημοσίευσης/συνδρομής όπου οι συνδρομές απαντώνται τόσο με υπάρχουσες όσο και με μελλοντικές δημοσιεύσεις.

Τα προβλήματα, τα οποία συνήθως μελετώνται σε ένα τέτοιο σύστημα αφορούν το φόρτο επικοινωνίας που επισύρουν τα εκάστοτε πρωτόκολλα, τόσο σε αριθμό μηνυμάτων ανακάλυψης όσο και μηνυμάτων περιεχομένου. Οι δύο αυτές παράμετροι αφορούν την ικανότητα κλιμάκωσης του συστήματος. Το προτεινόμενο πρωτόκολλο, όπως αποδεικνύεται στο σχετικό κεφάλαιο, βελτιώνει τη συμπεριφορά των επιδημικών αλγορίθμων εκμεταλλευόμενο το φαινόμενο της επαναληπτικότητας των ερωτημάτων των χρηστών.

Στο κομμάτι της ανακάλυψης του κατανεμημένου περιεχομένου και, ιδιαίτερα στα ομότιμα δίκτυα, όπου εστιάζει η παρούσα διατριβή, δύο είναι οι κυρίαρχες τάσεις: οι κατανεμημένοι πίνακες κατακερματισμού και τα επιδημικά πρωτόκολλα. Οι αλγόριθμοι κάθε μίας από αυτές τις κατηγορίες εφαρμόζονται ανάλογα με τις απαιτήσεις που προκύπτουν τόσο από τη μορφή των ερωτημάτων που υποβάλλονται στο σύστημα, όσο και από το ευρύτερο περιβάλλον μέσα στο οποίο λειτουργεί (π.χ. δυναμικά και χαμηλής συνδεσιμότητας συστήματα ευνοούν εν γένει τη χρήση επιδημικών αλγορίθμων).

Οι κατανεμημένοι πίνακες κατακερματισμού γενικά προτιμώνται όταν μπορεί να γίνει αναζήτηση του περιεχομένου με βάση κάποιο μοναδικό αναγνωριστικό, κλειδί. Τα προτεινόμενα πρωτόκολλα στη βιβλιογραφία, τα οποία και παρουσιάζονται στις αντίστοιχες ενότητες, παρουσιάζουν πολύ καλές ιδιότητες κλιμάκωσης, με τη χειρότερη περίπτωση να είναι λογαριθμικής τάξης ως προς το μέγεθος του δικτύου.

Τα επιδημικά πρωτόκολλα από την άλλη προτιμώνται αφενός σε πιο ευαίσθητα δίκτυα και αφετέρου όταν το ταίριασμα ενός ερωτήματος αναζήτησης είναι καλύτερα να γίνεται σε κάθε κόμβο περιεχομένου ξεχωριστά. Το τελευταίο μπορεί να χρειάζεται είτε λόγω κάποιας καλύτερης τεχνολογίας που μπορεί να διαθέτει ο εκάστοτε κόμβος είτε γιατί γίνεται πολύ ακριβό να ακολουθηθεί μια προσέγγιση κατανεμημένου πίνακα κατακερματισμού όπως παραπάνω. Τα επιδημικά πρωτόκολλα έχουν προβλήματα κλιμάκωσης στη γενική τους μορφή, αν και στη βιβλιογραφία έχουν προταθεί λύσεις, οι οποίες βελτιώνουν την απόδοσή τους σε συγκεκριμένα περιβάλλοντα.

1.3 Οργάνωση Περιεχομένου στον Τελικό Χρήστη

Στο επίπεδο του τελικού χρήστη, η πληροφορία περιλαμβάνεται γενικά σε αρχεία και, συνεπώς, η οργάνωσή της συνήθως συνίσταται στην οργάνωση αυτών των αρχείων σε καταλόγους. Έτσι, κάθε χρήστης, ανάλογα και με τις ανάγκες του δημιουργεί καταλόγους, στους οποίους τοποθετεί τα αντίστοιχα αρχεία, ώστε να είναι στη συνέχεια σε θέση να

τα εντοπίσει πιο εύκολα, χωρίς να χρειάζεται να τα αναζητήσει σε ολόκληρο το σύστημα αρχείων του.

Ωστόσο, όταν μπαίνει και το κομμάτι του κατανεμημένου περιεχομένου στην εξίσωση της οργάνωσης του περιεχομένου, τότε η κλίμακα γίνεται πολύ μεγαλύτερη. Παραδοσιακά συστήματα διαμοιρασμού αρχείων, όπως το Network File System (NFS) προσαρτούν ολόκληρους απομακρυσμένους καταλόγους, μη δίνοντας πολλές ελευθερίες στο χρήστη πελάτη.

Η επιτυχής διαχείριση περιεχομένου, με στόχο την ευκολότερη ανακάλυψή του, γίνεται επιτυχώς και σε μεγάλη κλίμακα με χρήση σημασιολογικών τεχνολογιών. Χαρακτηριστικό τέτοιο παράδειγμα αποτελεί ο σημασιολογικός ιστός, ο οποίος κερδίζει συνεχώς έδαφος, λόγω των μεγάλων δυνατοτήτων που παρέχει τόσο σε χρήστες όσο και σε παρόχους.

Το κατανεμημένο σύστημα αρχείων EuterpeFS που παρουσιάζεται στα πλαίσια αυτής της διατριβής, φιλοδοξεί να εισάγει τη χρήση σημασιολογικών τεχνολογιών για τη διαχείριση των αρχείων. Συγκεκριμένα, το EuterpeFS θεωρεί μια εικονική οργάνωση των καταλόγων, με βάση μια οντολογία, την οποία τροφοδοτεί ο χρήστης, ενώ μπορεί να την αλλάξει σε πραγματικό χρόνο, αν επιθυμεί διαφορετική αναπαράσταση.

Επιπλέον, κάνοντας χρήση πρότυπων τεχνολογιών, το EuterpeFS παρέχει ενοποιημένους τρόπους για την αναπαράσταση πληροφοριών για τα αρχεία, οι οποίες μπορεί να χρησιμοποιηθούν τόσο από άλλες εφαρμογές, όσο και από άλλες υπηρεσίες διευκολύνοντας έλεγχο πρόσβαση ή ακόμα και την ίδια την ανακάλυψη μέσω διαφορετικών μηχανισμών. Έτσι, ενισχύεται η ετερογένεια, ενώ δίνεται η δυνατότητα στους χρήστες να επιλέξουν μεταξύ διαφορετικών υλοποιήσεων των υπηρεσιών που απαιτούνται, ανάλογα με τις ανάγκες τους.

1.4 Διάρθρωση Διατριβής

Λαμβάνοντας υπόψη τα παραπάνω, η παρούσα διατριβή ακολουθεί μια προσέγγιση από το χαμηλότερο προς το υψηλότερο επίπεδο, ως εξής: το Κεφάλαιο 2 παρουσιάζει τις βασικές έννοιες των πληροφοριοκεντρικών δικτύων, ενώ στη συνέχεια δίνονται λεπτομέρειες σχετικά με τις δύο προτεινόμενες αρχιτεκτονικές για αποτελεσματικότερη υποστήριξη ομότιμης επικοινωνίας και ροής βίντεο, συνοδευόμενες από αντίστοιχες πειραματικές εκτιμήσεις της επίδρασης που έχουν στο σύστημα. Στη συνέχεια, στο Κεφάλαιο 3 παρουσιάζεται το πρωτόκολλο δημοσίευσης/συνδρομής Caching Content-based Publish/Subscribe (CCPS) που υλοποιήθηκε για να βελτιώσει την απόδοση των επιδημικών αλγορίθμων σε συστήματα δημοσίευσης/συνδρομής βασισμένα στο περιεχόμενο. Κατόπιν, στο Κεφάλαιο 4 γίνεται αναφορά στα τμήματα και τις τεχνολογίες των προτύπων MPEG-21 και MPEG-M, που χρειάστηκαν για την υλοποίηση ιδιαίτερα του κατανεμημένου συστήματος αρχείων

EuterpeFS, το οποίο παρουσιάζεται στο Κεφάλαιο 5. Τέλος, στο Κεφάλαιο 6 αναφέρονται τα συμπεράσματα αυτής της διατριβής, καθώς και οι μελλοντικές κατευθύνσεις που προκύπτουν από τη σχετική έρευνα.

Κεφάλαιο 2

Παράδοση Περιεχομένου με Πληροφοριοκεντρικά Δίκτυα

2.1 Εισαγωγή

Τα πληροφοριοκεντρικά δίκτυα (Information Centric Networks - ICN) αποτελούν ένα νέο μοντέλο αρχιτεκτονικής του επιπέδου δικτύου, σύμφωνα με το οποίο το δίκτυο θα πρέπει να δρομολογεί τα πακέτα με βάση το περιεχόμενο που μεταφέρουν και όχι με βάση την τοποθεσία όπου πρέπει να μεταφερθεί αυτό το περιεχόμενο. Το πλεονέκτημα αυτής της αρχιτεκτονικής είναι ότι το δίκτυο, εφόσον γνωρίζει τι πληροφορία μεταφέρει, είναι σε θέση να τη διαχειριστεί με πιο αποδοτικό τρόπο, με θετική επίδραση από τον πάροχο υπηρεσιών δικτύου, μέχρι τον τελικό χρήστη.

Η έρευνα στην περιοχή των πληροφοριοκεντρικών δικτύων κερδίζει εξάλλου όλο και περισσότερο το ενδιαφέρον των ερευνητών, όπως φαίνεται από τον αριθμό των ερευνητικών έργων παγκοσμίως, τα οποία εστιάζουν στην εν λόγω περιοχή, όπως τα SAIL [1], Named-Data Networking [2], PURSUIT [3], COMET [4], CONVERGENCE [5] [6] [7] [8] και ANR-CONNECT [9]. Παράλληλα, έχει δημιουργηθεί και μια ομάδα εργασίας προτυποποίησης στο Internet Research Task Force (IRTF) [10]. Αν και η έρευνα σε αυτήν την περιοχή βρίσκεται ακόμα στην αρχή, ο αριθμός των χρηματοδοτούμενων έργων που έχουν πραγματοποιηθεί, καθώς και η ενασχόληση ερευνητών που έχουν διαδραματίσει σημαίνοντα ρόλο στην εξέλιξη των σύγχρονων δικτύων IP, όπως του Van Jacobson, ο οποίος θεωρείται ένας από τους πρωτεργάτες των πληροφοριοκεντρικών δικτύων, αποτελούν ενθαρρυντικά στοιχεία για τη μελλοντική εξέλιξη και τελική ενσωμάτωση αυτών των τεχνολογιών.

Για να μπορέσει το δίκτυο να δρομολογήσει τα πακέτα με βάση το περιεχόμενο, θα πρέπει το περιεχόμενο αυτό να αντιπροσωπεύεται από κάποιο όνομα. Το όνομα αυτό αποτελεί το στοιχείο κλειδί για την προώθηση των αιτήσεων και των αποκρίσεων προς τον κατάλληλο προορισμό. Έτσι, μια αίτηση πελάτη αντιπροσωπεύεται από κάποιο μήνυμα

ενδιαφέροντος (*interest message*), το οποίο προωθείται προς τον κόμβο εξυπηρετητή (*servicing node*), ο οποίος αποτελεί το συνδεδετικό κρίκο μεταξύ του κατόχου της πληροφορίας και του δικτύου.

Ένα μεγάλο πλεονέκτημα που προκύπτει από τη δυνατότητα του δικτύου να γνωρίζει τι περιεχόμενο περνάει από τους κόμβους είναι η χρήση κρυφής μνήμης, η οποία μπορεί να έχει πολλαπλά οφέλη τόσο σε χρόνους απόκρισης όσο και σε χρησιμοποιούμενο εύρος ζώνης. Από την άλλη, όλη αυτή η γνώση που αποκτά το δίκτυο πάνω στο περιεχόμενο εγείρει και ζητήματα ασφάλειας και ιδιωτικότητας, τα οποία αναφέρονται σε επόμενα τμήματα του κεφαλαίου.

2.2 Σχήματα Ονομάτων για Πληροφοριοκεντρικά Δίκτυα

2.2.1 Γενικά

Η θεμελιώδης αρχή ενός πληροφοριοκεντρικού δικτύου, τουλάχιστον ως προς τον προσανατολισμό που υπάρχει στην τρέχουσα έρευνα, είναι ότι η πληροφορία δρομολογείται με βάση κάποιο όνομα, το οποίο της δίνεται από το χρήστη που την καθιστά διαθέσιμη πάνω από το εν λόγω δίκτυο. Το όνομα αυτό δύναται να καταστήσει ελκυστικό το πληροφοριοκεντρικό δίκτυο για δύο βασικούς λόγους: αφενός διότι δίνεται η δυνατότητα στο δίκτυο να διακρίνει πακέτα που φέρουν το ίδιο όνομα και, συνεπώς, να μπορεί να εξυπηρετήσει κάποια από αυτά μέσω κρυφής μνήμης και, αφετέρου, διότι οι χρήστες θα μπορούσαν να ζητούν απευθείας από το δίκτυο την πληροφορία με το όνομα, χωρίς να παρεμβάλλεται κάποια σύνθετη εφαρμογή ανακάλυψης, με απευθείας εκμετάλλευση του δικτύου.

Ο πρώτος από τους προαναφερθέντες λόγους, η εκμετάλλευση δηλαδή του ονόματος από το ίδιο το δίκτυο, δεν απαιτεί την ύπαρξη ονομάτων τα οποία είναι ευανάγνωστα προς το χρήστη -αρκεί η ίδια πληροφορία να αναπαρίσταται πάντα από το ίδιο όνομα. Αντίθετα, για να εκμεταλλευτεί ο χρήστης αυτή τη νέα δυνατότητα του δικτύου, το όνομα θα πρέπει να του είναι ευανάγνωστο, πράγμα το οποίο δημιουργεί ακόμα περισσότερες προκλήσεις, αναφορικά με την κλιμάκωση, και την ασφάλεια.

Λαμβάνοντας υπόψη την αρχιτεκτονική ενός πληροφοριοκεντρικού δικτύου όπως το CCNx [11] (το οποίο είναι και το πιο διαδεδομένο), είναι προφανές ότι η πληροφορία μπορεί να φτάσει στον αιτούμενο χωρίς η αίτηση του τελευταίου να έχει φτάσει στον ίδιο τον κάτοχο του περιεχομένου. Αυτό μπορεί να συμβεί εξαιτίας δύο λειτουργιών του δικτύου: είτε η αίτηση μπορεί να εξυπηρετηθεί από κάποια κρυφή μνήμη είτε μπορεί να εκκρεμεί ήδη μια προηγούμενη αίτηση στη διαδρομή της τρέχουσας, με αποτέλεσμα μόνο η πρώτη να φτάσει στον πάροχο του περιεχομένου.

Για τους λόγους αυτούς, το όνομα θα πρέπει να μπορεί να πιστοποιήσει την πληροφορία που αντιπροσωπεύει, καθώς σε διαφορετική περίπτωση ο οποιοσδήποτε κακόβουλος

χρήστης θα μπορούσε να αντικαταστήσει την πραγματική πληροφορία που αντιπροσωπεύει το όνομα με ψευδή, πραγματοποιώντας επιθέσεις πλαστογράφησης (spoofing) ή άρνησης υπηρεσίας (denial of service). Τα λεγόμενα αυτοπιστοποιούμενα ονόματα θεωρείται ότι μπορούν να δώσουν απαντήσεις σε αυτά τα προβλήματα ασφάλειας, ενώ παράλληλα, μέσα από αυτή τη διαδικασία παρέχουν και δυνατότητα για βελτίωση της κλιμάκωσης [12].

2.2.2 Η προσέγγιση του TRIAD

Το Translating Relaying Internet Architecture integrating Active Directories (TRIAD) αποτελεί μια από τις παλαιότερες προσεγγίσεις στο χώρο των πληροφοριοκεντρικών δικτύων. Παρόλο που ως προσέγγιση δε χρησιμοποιείται πλέον ως βάση για έρευνα, η προσέγγιση που προτείνει ως προς το σχήμα ονομάτων και τη δρομολόγηση με βάση το όνομα συνεχίζει να συναντάται στη σύγχρονη βιβλιογραφία.

Η προσέγγιση του TRIAD [13] βασίζεται εν πολλοίς στο μοντέλο του DNS, όπου οι δρομολογητές, οι οποίοι πλέον ονομάζονται Δρομολογητές Περιεχομένου (Content Routers - CR) διατηρούν αντιστοίχιση ανάμεσα σε ονόματα και διευθύνσεις IP όπου βρίσκεται το περιεχόμενο. Τα πρωτόκολλα, στα οποία βασίζονται οι δρομολογητές περιεχομένου για να πραγματοποιήσουν τη δρομολόγηση με βάση το όνομα είναι το Internet Name Resolution Protocol (INRP) και το Name-Based Routing Protocol (NBRP).

2.2.2.1 Internet Name Resolution Protocol

Το INRP ακολουθεί παρόμοια λογικά με τη κλασική δρομολόγηση σε δίκτυα IP, με τη διαφορά όμως ότι σε αυτήν την περίπτωση οι πίνακες δρομολόγησης αντί να έχουν αντιστοίχιση μεταξύ διεύθυνσης IP και επόμενου δρομολογητή (next hop), περιλαμβάνουν αντιστοίχιση ονόματος (ή μέγιστου προθέματος) με τον επόμενο δρομολογητή.

Καθώς μπορεί να υπάρχει περιεχόμενο με επικάλυψη στο πρόθεμα του ονόματος, το οποίο βρίσκεται κατανεμημένο σε υπολογιστές που συνδέονται με διαφορετικούς δρομολογητές, είναι πιθανό σε κάποια σημεία της διαδρομής της αίτησης να υπάρχουν διάφοροι, ισοδύναμοι, πιθανοί προορισμοί. Έτσι, κάθε φορά επιλέγεται μια διαδρομή, η οποία, όμως, μπορεί να επιφέρει και το καλύτερο δυνατό αποτέλεσμα.

Αν και το INRP δεν το ορίζει αυστηρά, παρόλα αυτά αναφέρεται ως δυνατότητα, η επιλογή ενός εναλλακτικού μονοπατιού σε κάποιον ενδιάμεσο δρομολογητή, σε περίπτωση που δεν εντοπιστεί αποτέλεσμα στο προηγούμενο μονοπάτι. Επίσης, παρέχει τη δυνατότητα στην εφαρμογή πελάτη να αποκλείει παρόχους περιεχομένου, οι οποίοι π.χ. εμφανίζουν χαμηλή διαθεσιμότητα. Δίνεται, δηλαδή, η δυνατότητα anycast την οποία μπορούν να εκμεταλλευτούν ποικιλοτρόπως οι εφαρμογές, ώστε να βελτιστοποιήσουν την τελική εμπειρία χρήστη.

Αξίζει να σημειωθεί ότι το INRP προωθεί τη χρήση των ονομάτων των εξυπηρε-

τητών (βλ. HTTP URL) για την εύρεση της διαδρομής, το οποίο αποτελεί προπομπό για τα συστήματα ονομάτων, τα οποία προτείνονται πλέον στη βιβλιογραφία, σύμφωνα με τα οποία βοηθά στην κλιμάκωση η δρομολόγηση με βάση κάποιο αναγνωριστικό χρήστη, παρά με ολόκληρο το όνομα. Αυτό το σχήμα, άλλωστε, ακολουθήθηκε και από το DONA που θα παρουσιαστεί σε επόμενη ενότητα, με τη διαφορά ότι εκεί χρησιμοποιείται το αναγνωριστικό του χρήστη

2.2.2.2 Name-Based Routing Protocol

Το NBRP χρησιμοποιείται σε αντιστοιχία με το BGP [14] για να ενισχύσει τη σύνδεση μεταξύ των δρομολογητών που βρίσκονται σε διαφορετικά αυτόνομα συστήματα. Έτσι, όπως το BGP παρέχει πληροφορία για το πρόθεμα των διευθύνσεων IP και το πώς αυτές είναι καταναμημένες στα διάφορα αυτόνομα συστήματα, δίνοντας επιπλέον και τις απαραίτητες οδηγίες στα πακέτα για το πώς θα φτάσουν σε αυτά τα συστήματα, έτσι και το NBRP παρέχει αντίστοιχες πληροφορίες για τα προθέματα ονομάτων του δημοσιευμένου περιεχομένου στα εκάστοτε αυτόνομα συστήματα.

Συγκεκριμένα, όταν δημοσιεύεται νέο περιεχόμενο, τότε η εγγραφή που δημιουργείται στον εκάστοτε δρομολογητή περιεχομένου θα πρέπει να περιλαμβάνει τη διεύθυνση όπου θα πρέπει να ακολουθήσει το πακέτο για το επόμενο βήμα καθώς και το μονοπάτι από δρομολογητές που θα πρέπει να ακολουθηθεί προκειμένου να φτάσει κάποια αίτηση για αυτό το περιεχόμενο στον υπολογιστή που το παρέχει (κόμβος -εξυπηρετητής).

Ένα επιπλέον χαρακτηριστικό του NBRP είναι ότι είναι σε θέση να επιβεβαιώνει πρώτα την αυθεντικότητα του περιεχομένου, προτού αποφασίσει να προωθήσει τη δημοσίευσή του. Έτσι, ο κόμβος-εξυπηρετητής θεωρείται ότι υπογράφει ψηφιακά το μήνυμα δημοσίευσης του περιεχομένου που στέλνεται στους εκάστοτε δρομολογητές, ενώ έχει επιπλέον και τη δυνατότητα να καθορίσει ρητά αν κάποιος δρομολογητής έχει δικαίωμα να προωθεί αιτήσεις για το εν λόγω περιεχόμενο.

2.2.3 Η προσέγγιση του DONA

Το Data-Oriented Network Architecture (DONA) [15] αποτελεί μια από τις πιο διαδεδομένες σύγχρονες λύσεις στο πρόβλημα της ονοματολογίας των πληροφοριοκεντρικών δικτύων. Ο σχεδιασμός του έχει γίνει με τέτοιο τρόπο, ώστε να διαχωρίζονται οι λειτουργίες της ονομασίας αυτής καθεαυτής από τη λειτουργία της ανεύρεσης του περιεχομένου με βάση το όνομα. Επιπλέον, το DONA ορίζει τρεις βασικούς άξονες που διέπουν τη δομή του: τη διατήρηση, τη διαθεσιμότητα και την αυθεντικότητα των ονομάτων, καθώς και του περιεχομένου που αυτά αντιπροσωπεύουν.

Από τους προαναφερθέντες άξονες, το DONA έχει σχεδιαστεί με τέτοιο τρόπο ώστε οι δύο πρώτοι να ικανοποιούνται από τον τρόπο που έχει σχεδιαστεί το σχήμα των ονομά-

των, ενώ ο τρίτος ικανοποιείται από την υπηρεσία που πραγματοποιεί την ανεύρεση της αντιστοίχησης του ονόματος με τον κόμβο-εξυπηρετητή του περιεχομένου.

Οι ιδιότητες της διατήρησης και της αυθεντικότητας των ονομάτων ικανοποιούνται με τη χρήση επίπεδων ονομάτων, τα οποία έχουν την επιπλέον ικανότητα να μπορούν να αυθεντικοποιηθούν από μόνα τους, χωρίς κάποια επιπρόσθετη πληροφορία, όπως θα παρυσιαστεί σε επόμενη ενότητα. Το σχήμα των ονομάτων που θα προταθεί, μπορεί να μην είναι εύχρηστο καθώς περιλαμβάνει μεγάλο αριθμό από bits, τα οποία είναι αδύνατο να θυμάται κάποιος χρήστης, αλλά από την άλλη, ακόμα και σήμερα ούτε τις διευθύνσεις IP χρειάζεται να τις θυμάται ο χρήστης -απεναντίας, για το σκοπό αυτό χρησιμοποιούνται εξωτερικές υπηρεσίες ευρετηρίων. Η ιδιότητα της διαθεσιμότητας ικανοποιείται με την παροχή του πλησιέστερου δυνατού αντιγράφου του περιεχομένου που αιτείται κάποιος πελάτης (όπως παρουσιάστηκε σε προηγούμενη ενότητα, το πληροφοριοκεντρικό δίκτυο έχει την ικανότητα να κρατά αντίγραφα του περιεχομένου σε κρυφές μνήμες) και την αποφυγή εξυπηρετητών που είναι ήδη πολύ απασχολημένοι ή έχουν βγει εκτός σύνδεσης. Όλα αυτά θα αναλυθούν με λεπτομέρειες στις επόμενες υποενότητες.

2.2.3.1 Σχήμα Ονομάτων του DONA

Το σχήμα ονομάτων του DONA βασίζεται στην πληροφορία για την οντότητα κάτοχο ή πάροχο του περιεχομένου. Συγκεκριμένα, κάθε τέτοια οντότητα αντιπροσωπεύεται από ένα ζεύγος κλειδιών, δημόσιο και ιδιωτικό, τα οποία χρησιμοποιούνται για να εξασφαλίσουν τόσο τη μοναδικότητα των ονομάτων όσο και την αυθεντικότητά τους.

Συγκεκριμένα, κάθε όνομα πρέπει να φέρει οπωσδήποτε ένα αναγνωριστικό, το οποίο αντιστοιχίζεται με μοναδικό τρόπο στην οντότητα που παρέχει το περιεχόμενο. Αυτό εξασφαλίζεται θεωρώντας σε κάθε όνομα ένα πρόθεμα P (από τη λέξη principal), το οποίο είναι το αποτέλεσμα της εφαρμογής μιας κρυπτογραφικής συνάρτησης κατακερματισμού (hash) πάνω στο δημόσιο κλειδί του παρόχου. Στη συνέχεια, το όνομα περιλαμβάνει και το επίθεμα που είναι η λεγόμενη ετικέτα (label) L, και η χρήση της είναι στην αποκλειστική ευθύνη του χρήστη, ο οποίος είναι και υπεύθυνος για τη χρήση μοναδικών τέτοιων ετικετών.

Τα ονόματα P:L του DONA εξασφαλίζουν ότι κάθε χρήστης είναι υπεύθυνος για τη μοναδικότητα του ονόματος του περιεχομένου, ενώ η δρομολόγηση γίνεται με βάση μόνο το P, αφού είναι με πολύ μεγάλη πιθανότητα μοναδικό (όπως προκύπτει από τους αντίστοιχους κρυπτογραφικούς αλγορίθμους). Έτσι, όταν η αίτηση φτάσει στον κόμβο-εξυπηρετητή που αντιστοιχεί στην εν λόγω οντότητα, ο τελευταίος θα αναζητήσει το περιεχόμενο, το οποίο φέρει το κομμάτι της ετικέτας στο όνομα.

Το DONA θεωρεί ότι, εκτός από το όνομα, τα πακέτα πρέπει επίσης να φέρουν το δημόσιο κλειδί καθώς και την ψηφιακή υπογραφή των δεδομένων που μεταφέρουν. Η υπογραφή αυτή δύναται να επιβεβαιωθεί τόσο από κάθε κόμβο του δικτύου, όσο και από τον

τελικό αποδέκτη του πακέτου, καθώς προκύπτει κάνοντας χρήση του ιδιωτικού κλειδιού που αντιστοιχεί στο δημόσιο που φέρει το πακέτο. Με αυτόν τον τρόπο εξασφαλίζεται η ιδιότητα της αυθεντικότητας των δεδομένων που μεταφέρονται πάνω από το πληροφοριοκεντρικό δίκτυο.

Η διατήρηση των ονομάτων επιβεβαιώνεται από το γεγονός ότι τα ονόματα δεν αναφέρονται σε μια συγκεκριμένη τοποθεσία με την οποία είναι συνδεδεμένο το περιεχόμενο που αντιπροσωπεύουν -απεναντίας με τη χρήση τους, η πληροφορία δύναται να είναι αποθηκευμένη οπουδήποτε στο δίκτυο. Με αυτόν τον τρόπο, όσο τα δεδομένα υφίστανται, το όνομα και, άρα, και το περιεχόμενο αυτό καθ'αυτό θα διατηρείται.

2.2.3.2 Επίλυση Ονομάτων στο DONA

Η επίλυση ονομάτων στο DONA δε διαφέρει κατά πολύ από το TRIAD ή οποιοδήποτε άλλο αντίστοιχο σύστημα, όπως για παράδειγμα το DNS. Έτσι, και σε αυτήν την περίπτωση, κάθε εξυπηρετητής περιεχομένου θα πρέπει να δηλώσει σε κάποια τρίτη οντότητα ότι παρέχει πληροφορία υπό κάποιο συγκεκριμένο όνομα. Αντίστοιχα, σε αυτήν την οντότητα ανατρέχει και ο εκάστοτε πελάτης, ο οποίος πραγματοποιεί μια αίτηση για το περιεχόμενο που αντιστοιχεί σε αυτό το όνομα.

Όπως αναλύθηκε σε προηγούμενη ενότητα, στη βιβλιογραφία αναφέρονται δύο πιθανοί τρόποι για να γίνει η δρομολόγηση σε ένα πληροφοριοκεντρικό δίκτυο με βάση το όνομα του περιεχομένου: η αναζήτηση με βάση το όνομα σε μια κατανεμημένη βάση δεδομένων, η οποία συνίσταται από τους δρομολογητές του δικτύου και η δρομολόγηση με βάση το όνομα στην κοντινότερη τοποθεσία όπου βρίσκεται σχετικό αντίγραφο του δεδομένου. Στην πρώτη περίπτωση έχουμε συνήθως σταθερούς χρόνους αναζήτησης (π.χ. με τους δρομολογητές οργανωμένους σε ένα DHT, ο χρόνος αυτός είναι της τάξης $O(\log N)$, όπου N ο αριθμός των δρομολογητών του δικτύου) με την κάθε αναζήτηση να επιστρέφει τη διεύθυνση IP προορισμού. Αντίθετα, στη δεύτερη περίπτωση, πραγματοποιείται δρομολόγηση ακολουθώντας το συντομότερο μονοπάτι προς ένα γνήσιο αντίγραφο του περιεχομένου που αναζητείται. Το DONA έχει σχεδιαστεί ακολουθώντας τη δεύτερη περίπτωση της δρομολόγησης με βάση το όνομα.

Η οντότητα που ορίζει το DONA ως υπεύθυνη για την επίλυση των ονομάτων είναι ο χειριστής επίλυσης (resolution handler - RH). Στην πραγματικότητα, βέβαια, δε μιλάμε για μια τέτοια οντότητα, αλλά για ένα κατανεμημένο σύστημα από RH, οι οποίες είναι σε θέση να συνεργάζονται μεταξύ τους και να χειρίζονται τόσο τις δημοσιεύσεις όσο και τις αιτήσεις για περιεχόμενο.

Για να χειριστεί αυτές τις δύο περιπτώσεις, το DONA ορίζει δύο θεμελιώδεις τύπους μηνυμάτων: το REGISTER(P:L) και το FIND(P:L). Ο πρώτος τύπος μηνυμάτων χρησιμοποιείται από τον πάροχο του εκάστοτε περιεχομένου για τη δημοσίευσή του στο δίκτυο των RH ώστε αυτό να καταστεί διαθέσιμο στους πελάτες. Ο δεύτερος τύπος μηνυμάτων προέρχε-

ται από τους πελάτες και είναι ουσιαστικά η αίτηση που πραγματοποιούν οι τελευταίοι προκειμένου να προσπελάσουν το περιεχόμενο που φέρει το όνομα που δίνεται στην αίτηση.

Η αρχιτεκτονική των RH είναι ιεραρχική καθώς δύναται να υπάρχουν RH σε διάφορα επίπεδα, ξεκινώντας από αυτό του οργανισμού που ανήκει ο χρήστης, της εταιρείας, του παρόχου υπηρεσιών διαδικτύου κλπ ως το επίπεδο αυτόνομων συστημάτων. Η RH όπου αναφέρεται κάθε πελάτης για να δημοσιεύσει ή να ζητήσει περιεχόμενο είναι αποτυπωμένη σε κάποια ρύθμιση του συστήματός του, κατά αντιστοιχία με αυτό που συμβαίνει σήμερα για το DNS.

Όταν κάποιος κόμβος αποφασίσει να δημοσιεύσει το περιεχόμενό του, τότε στέλνει ένα μήνυμα REGISTER(P:L) στην RH του. Αυτό το μήνυμα REGISTER προωθείται προς τα ανώτερα επίπεδα της ιεραρχίας των RH μέχρι: (α) να βρεθεί εγγραφή με το ίδιο όνομα, η οποία βρίσκεται κοντινότερα από την τρέχουσα ή (β) να εμποδίσει τη διάδοση του μηνύματος κάποιος κανόνας/πολιτική που ισχύει σε ανώτερο επίπεδο και έχει να κάνει με τη σύνδεση αυτόνομων συστημάτων ή τη χρήση firewall για κάποια συγκεκριμένα πακέτα κλπ.

Κάθε φορά που λαμβάνεται ένα μήνυμα REGISTER από μία RH, θα πρέπει να επιβεβαιώνεται η αυθεντικότητά του. Στο πρώτο επίπεδο της τοπικής RH, η οποία λαμβάνει το μήνυμα REGISTER από τον πάροχο του περιεχομένου, αυτό γίνεται με επιβεβαίωση του P (από το P:L του ονόματος). Για να επιβεβαιώσει το P, η RH επιστρατεύει ένα σχήμα πρόκλησης-απόκρισης, όπου στέλνει ένα τυχαίο αλφαριθμητικό στον πάροχο, αυτός το στέλνει πίσω κρυπτογραφημένο με το ιδιωτικό του κλειδί, οπότε η RH το αποκρυπτογραφεί με το δημόσιό του κλειδί, το οποίο είτε προκύπτει απευθείας από το P είτε στέλνεται ξεχωριστά σαν ψηφιακό πιστοποιητικό, το οποίο επιβεβαιώνει και τη σύνδεση του P με τον εν λόγω πάροχο. Αν το αλφαριθμητικό που προέκυψε από την αποκρυπτογράφηση είναι το ίδιο με το αρχικό που στάλθηκε, τότε το REGISTER επιβεβαιώνεται και προωθείται στα ανώτερα στρώματα της ιεραρχίας των RH. Για να αποφευχθεί αυτή η διαδικασία σε κάθε βήμα, η πρώτη RH υπογράφει ψηφιακά το μήνυμα REGISTER που προωθεί, οπότε από εκεί και πέρα η διαδικασία προχωρά με βάση τη (δεδομένη) σχέση εμπιστοσύνης μεταξύ των RH.

Όπως έχει ήδη αναφερθεί σε προηγούμενη ενότητα, το σχήμα ονομάτων του DONA παρέχει τη δυνατότητα δρομολόγησης μόνο με βάση το P κομμάτι του ονόματος. Έτσι, ο πάροχος του περιεχομένου αρκεί να δηλώσει στο REGISTER μήνυμα ότι αυτός παρέχει το περιεχόμενο με το εν λόγω πρόθεμα, με αποτέλεσμα όλες οι σχετικές αιτήσεις να προωθούνται προς αυτόν. Φυσικά αυτή η δυνατότητα δεν αποκλείει τη δημοσίευση περιεχομένου με βάση και την ετικέτα του. Σε περίπτωση που μια RH έχει στους πίνακες δρομολόγησης ολόκληρο το όνομα P:L, τότε πραγματοποιεί προώθηση του πακέτου με βάση το μέγιστο ταίριασμα προθέματος.

Έχοντας αναλύσει τον τρόπο με τον οποίο λειτουργεί η διαδικασία της δημοσίευ-

σης του περιεχομένου, η διαδικασία εντοπισμού του προκύπτει σε πλήρη αντιστοιχία. Έτσι, όταν ένας πελάτης στέλνει στο δίκτυο ένα μήνυμα FIND(P:L), το μήνυμα αυτό το παίρνει αρχικά η τοπική RH, η οποία επιλέγεται με τον ίδιο τρόπο όπως και για το REGISTER. Στη συνέχεια, ελέγχει τον πίνακα δρομολόγησης και προωθεί το πακέτο είτε στην RH που για την οποία υπάρχει σχετική εγγραφή, είτε στην RH που βρίσκεται στο ακριβώς ανώτερο επίπεδο ιεραρχίας, καθώς η τελευταία έχει ευρύτερη αντίληψη περί του διαθέσιμου περιεχομένου.

2.3 Αρχιτεκτονικές Πληροφοριοκεντρικών Δικτύων

Όπως αναφέρθηκε και στην Εισαγωγή του Κεφαλαίου, η κεντρική ιδέα που βρίσκεται πίσω από τα πληροφοριοκεντρικά δίκτυα είναι ότι τα πακέτα δε δρομολογούνται με βάση μια διεύθυνση προορισμού, αλλά ένα όνομα. Έτσι, ακολουθώντας το μοντέλο publish-subscribe [16], οι προτεινόμενες αρχιτεκτονικές πληροφοριοκεντρικών δικτύων παρέχουν, αφενός, τη δυνατότητα στον κάτοχο περιεχομένου να δημοσιεύσει αυτό το περιεχόμενο (publish) στο δίκτυο με κάποιο όνομα, το οποίο δύναται να είναι αναγνώσιμο από το χρήστη (υπάρχουν διάφορα σχήματα ονομάτων, αναγνώσιμα ή μη, που αναφέρονται σε επόμενη ενότητα), και, αφετέρου, τη δυνατότητα σε κάποιον ενδιαφερόμενο να αναζητήσει (lookup/subscribe) το περιεχόμενο που επιθυμεί με το ίδιο όνομα.

2.3.1 Μοντέλο Υπερκείμενου Δικτύου

Ένα βασικό σημείο διαφοροποίησης των προτεινόμενων αρχιτεκτονικών είναι ότι άλλες υποστηρίζουν και άλλες όχι τη λειτουργία subscribe. Έτσι, η αρχιτεκτονική του i3 [17], βασίζεται στην ιδέα του κατανεμημένου πίνακα κατακερματισμού (Distributed Hash Table - DHT) και πιο συγκεκριμένα σε προηγούμενη δουλειά που συμμετείχαν μέλη της εν λόγω ομάδας, το Chord [18], υποστηρίζει τη λειτουργία subscribe κάνοντας χρήση των λεγόμενων *κόμβων συνάντησης* (*rendez-vous nodes*).

Η αρχιτεκτονική προσέγγιση αυτής της ομάδας πρόκειται ουσιαστικά για μια προσέγγιση υπερδικτύου, το οποίο δημιουργείται από τους κόμβους εξυπηρετητές του δικτύου, το οποίο κατά τα άλλα διατηρεί την κλασική αρχιτεκτονική που υπαγορεύει το πρωτόκολλο IP. Οι κόμβοι συνάντησης διατηρούν έναν κατανεμημένο πίνακα κατακερματισμού με τον οποίο αντιστοιχίζουν τα ονόματα των μηνυμάτων σε διευθύνσεις των κόμβων που παρέχουν το περιεχόμενο. Έτσι, κάθε μήνυμα έχει τη μορφή $\langle id, addr \rangle$, όπου *id* είναι το αναγνωριστικό (όνομα) και *addr* ο συνδυασμός διεύθυνσης IP και πόρτας του εξυπηρετητή.

Ένα πλεονέκτημα αυτού του σχήματος είναι η εγγενής υποστήριξη κινητικότητας (mobility), πολυεκπομπής (multicast) και επιλογή βέλτιστης από πολλαπλές διαδρομές (anycast). Λεπτομερέστερα, για την υποστήριξη της κινητικότητας, το i3 παρέχει τη δυνατότητα αλλαγής του *addr* σε ένα ζεύγος $\langle id, addr \rangle$ με κάποιο νεότερο. Έτσι, όταν κάποιος

κόμβος αλλάζει θέση (άρα και διεύθυνση), ανανεώνει τη σχετική εγγραφή στον κόμβο συνάντησης, ο οποίος πλέον δείχνει στη νέα διεύθυνση. Η πολυεκπομπή επιτυγχάνεται με έναν επίσης απλό τρόπο. Συγκεκριμένα, στους κόμβους συνάντησης είναι δυνατό να υπάρχει 1:N αντιστοίχιση ανάμεσα σε όνομα και διεύθυνση. Με αυτόν τον τρόπο, δημιουργείται μια ομάδα πολυεκπομπής στον εν λόγω κόμβο συνάντησης, οπότε όταν γίνει αίτηση για ένα τέτοιο αναγνωριστικό, επιστρέφονται όλες οι διευθύνσεις των κόμβων που ανταποκρίνονται στο αντίστοιχο περιεχόμενο. Τέλος, για τη λειτουργία της επιλογής ενός μόνο κόμβου από την ομάδα (anycast), το i3 παρέχει τη δυνατότητα του μέγιστου ταιριάσματος αναγνωριστικού. Συγκεκριμένα, οι κόμβοι που ανήκουν σε μια ομάδα anycast περιλαμβάνουν στις διευθύνσεις τους ένα σύνολο από k bits τα οποία θα πρέπει να ταιριάξουν με αντίστοιχα m bits του πακέτου της αίτησης. Η αίτηση θα δρομολογηθεί προς εκείνο τον κόμβο που έχει τα περισσότερα κοινά bits, ξεκινώντας από το πρώτο.

Παρόλο που αυτή η αρχιτεκτονική φαντάζει ιδιαίτερα ελκυστική, καθώς παρέχει τη δυνατότητα δημιουργίας μιας υποδομής πληροφοριοκεντρικού δικτύου χωρίς καμία αλλαγή στην υφιστάμενη αρχιτεκτονική και λειτουργία των υπαρχόντων πρωτοκόλλων, υπάρχουν σημεία που δημιουργούν σκεπτικισμό. Έτσι, αυτό το αρχιτεκτονικό μοντέλο δεν αποτελεί μια ξεκάθαρη, παρά μια υβριδική, προσέγγιση στο θέμα των πληροφοριοκεντρικών δικτύων και φαντάζει περιοριστική ως προς τις νέες δυνατότητες που θα μπορούσε να ανοίγει μια καθαρή (clean slate) προσέγγιση. Επιπλέον, η δρομολόγηση με χρήση κατανεμημένου πίνακα κατακερματισμού προσφέρει μεν το άνω όριο $O(\log N)$ στα μηνύματα που χρειάζονται που σταλούν προκειμένου να εντοπιστεί το περιεχόμενο, αλλά από την άλλη δεν επιτρέπει τη χρήση αλγορίθμων εύρεσης συντομότερου μονοπατιού καθώς και άλλων τεχνικών βελτιστοποίησης που θα μπορούσαν να εφαρμοστούν κατά αναλογία με τα τυπικά IP δίκτυα.

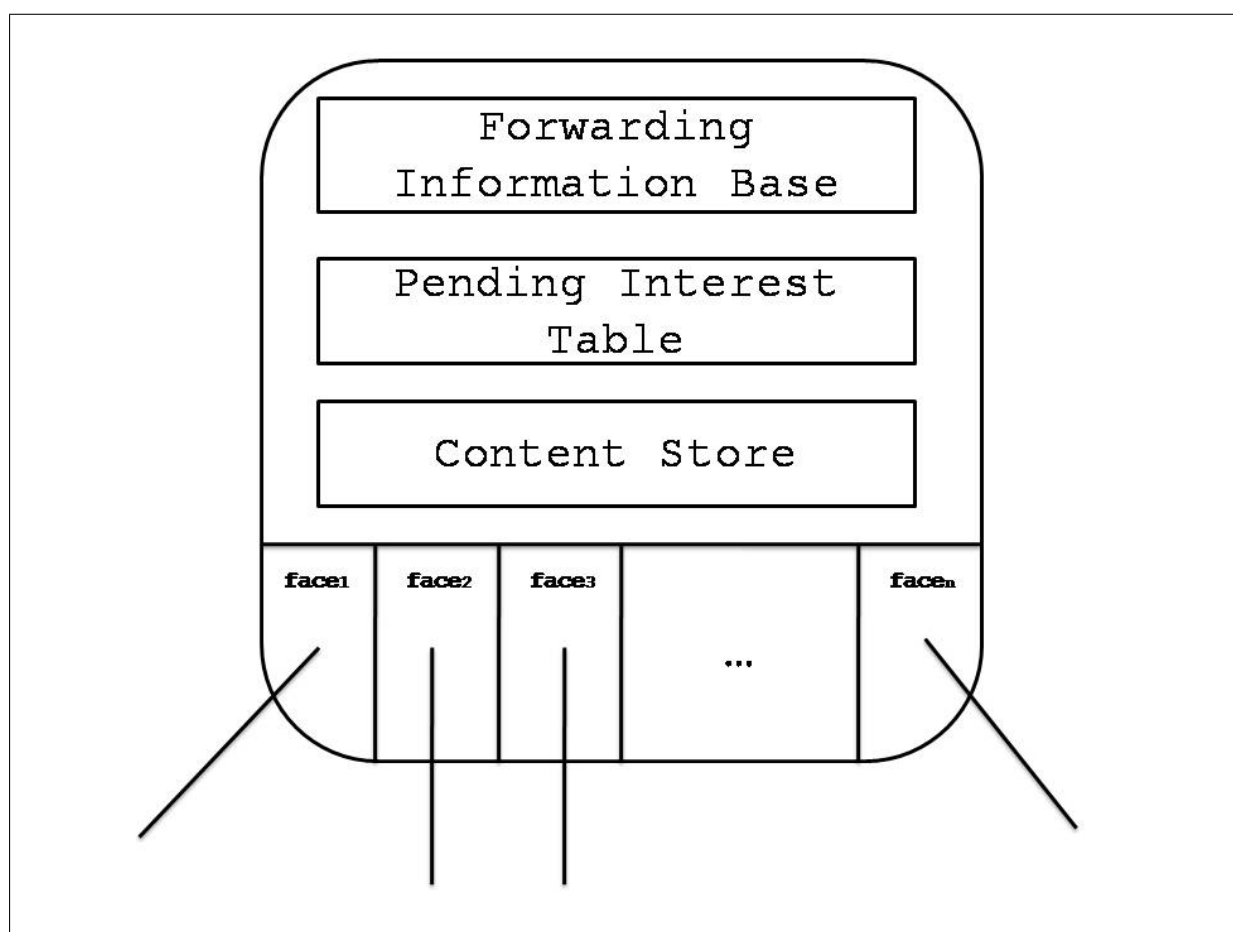
2.3.2 Μοντέλο Καθαρού Μητρώου

Το αρχιτεκτονικό μοντέλο καθαρού μητρώου (clean slate) ορίζει τη δημιουργία μιας αρχιτεκτονικής, η οποία έχει σχεδιαστεί ειδικά για ένα πληροφοριοκεντρικό δίκτυο με όλα τα ιδιαίτερα χαρακτηριστικά του. Ένα βασικό θέμα που έχει να κάνει με την οργάνωση ενός τέτοιου δικτύου είναι το σχήμα των ονομάτων των πακέτων, καθώς αυτό συνδέεται με το μέγεθος των πινάκων δρομολόγησης, το οποίο θα εξεταστεί σε επόμενη ενότητα.

Μία από τις πιο χαρακτηριστικές δουλειές στο χώρο των πληροφοριοκεντρικών δικτύων αποτελεί το έργο CCNx [11] που εκπονείται κατά κύριο λόγο από το Palo Alto Research Center (PARC) και το οποίο κάνει μια προσέγγιση καθαρού μητρώου στην αρχιτεκτονική [19]. Το μοντέλο που ορίζει το CCNx έχει χρησιμοποιηθεί και από άλλες προσεγγίσεις, όπως το CONET [20], καθώς προβλέπει μια καθαρή αρχιτεκτονική. Επιπλέον, το CCNx θεωρείται και το μοντέλο υποκείμενου δικτύου στην παρούσα διατριβή, όπως θα διαφανεί σε επόμενη ενότητα του κεφαλαίου.

2.3.2.1 Αρχιτεκτονική Κόμβου CCNx

Η αρχιτεκτονική ενός κόμβου CCNx, όπως φαίνεται και στην Εικόνα 1, αποτελείται από τέσσερα βασικά στοιχεία: τις διεπαφές του δικτύου (*faces*), την αποθήκη περιεχομένου (Content Store - CS), τον πίνακα των ενδιαφερόντων που δεν έχουν ικανοποιηθεί ακόμα (Pending Interest Table - PIT) και τον πίνακα προώθησης της πληροφορίας (Forwarding Information Base - FIB), ο οποίος ουσιαστικά αποτελεί τον πίνακα δρομολόγησης προς την πληροφορία.



Σχήμα 1: Η Αρχιτεκτονική ενός Κόμβου CCNx

Η αποθήκη περιεχομένου (CS) είναι ουσιαστικά η αλλιώς λεγόμενη *κρυφή μνήμη* (*cache*) στην οποία αποθηκεύονται τα μηνύματα περιεχομένου, τα οποία διέρχονται μέσα από το πληροφοριοκεντρικό δίκτυο. Το αρχιτεκτονικό μοντέλο δεν ορίζει το μέγεθος που μπορεί να έχει αυτή η κρυφή μνήμη, ούτε αλγορίθμους αντικατάστασης, αν και όσον αφορά τους τελευταίους γενικά προτιμώνται αυτοί που μεγιστοποιούν την πιθανότητα επαναχρησιμοποίησης, όπως οι Least Recently Used (LRU) ή Least Frequently Used (LFU).

Ο πίνακας εκκρεμών ενδιαφερόντων (PIT) αποτελεί έναν ακόμα τρόπο αποφόρτισης του δικτύου από μηνύματα, εκμεταλλευόμενο το γεγονός ότι γνωρίζει για ποιο ακρι-

βώς περιεχόμενο έχει εκφραστεί το ενδιαφέρον. Έτσι, όπως θα εξηγηθεί και παρακάτω, ο εν λόγω πίνακας περιλαμβάνει εγγραφή για κάθε μήνυμα ενδιαφέροντος που έχει προωθηθεί και δεν έχει ληφθεί ακόμα απάντηση, ώστε, αν περάσει και νέο μήνυμα στο μεταξύ, να μην προωθηθεί και αυτό προς τον ίδιο προορισμό με το προηγούμενο για οικονομία χωρητικότητας του δικτύου.

Ο πίνακας προώθησης περιεχομένου (FIB) είναι ο πίνακας που περιλαμβάνει την απαραίτητη πληροφορία για να προωθηθούν τα μηνύματα ενδιαφέροντος προς τον τελικό τους προορισμό, όπου βρίσκεται το περιεχόμενο. Πρόκειται, δηλαδή, για μια παραλλαγή ενός κλασικού πίνακα δρομολόγησης, μόνο που αντί για διευθύνσεις, περιλαμβάνει ονόματα (ή προθέματα, ανάλογα με τον αλγόριθμο) και η αντιστοίχιση γίνεται ανάμεσα σε όνομα και face, με το τελευταίο να αντιστοιχίζεται συνήθως σε κάποια σύνδεση με κάποιον άλλο κόμβο (next hop).

2.3.2.2 Βασικοί Τύποι Μηνυμάτων CCNx

Το CCNx θεωρεί δύο θεμελιώδεις τύπους μηνυμάτων: το μήνυμα ενδιαφέροντος (Interest Message) και το μήνυμα περιεχομένου (Content Message/Object). Το μήνυμα ενδιαφέροντος φέρει το ενδιαφέρον του χρήστη για πληροφορία υπό συγκεκριμένο όνομα, το οποίο περιλαμβάνεται στο μήνυμα. Όταν ένα μήνυμα ενδιαφέροντος φτάσει σε κάποιο κόμβο του δικτύου, τότε ο κόμβος αυτός αρχικά θα κοιτάξει την κρυφή μνήμη του (CS) για την περίπτωση που η αιτούμενη πληροφορία βρίσκεται εκεί αποθηκευμένη, οπότε η τελευταία προωθείται στον κόμβο από τον οποίο παρελήφθη το μήνυμα ενδιαφέροντος. Σε αντίθετη περίπτωση, ελέγχεται ο πίνακας PIT για την περίπτωση που έχει γίνει κάποια άλλη αίτηση για το ίδιο περιεχόμενο, η οποία εκκρεμεί, οπότε και η διαδρομή του μηνύματος ενδιαφέροντος σταματά σε αυτόν τον κόμβο. Αν δε βρεθεί η πληροφορία ούτε στο PIT, τότε ελέγχεται ο πίνακας δρομολόγησης FIB, ο οποίος θα πρέπει να περιέχει τον επόμενο κόμβο, ο οποίος έχει πληροφορία για το πού βρίσκεται η αιτούμενη πληροφορία και προωθεί το μήνυμα από το αντίστοιχο face. Σε περίπτωση που δεν υπάρχει σχετική εγγραφή στο FIB, το μήνυμα ενδιαφέροντος απορρίπτεται.

Το μήνυμα περιεχομένου φέρει την πληροφορία αυτή καθαυτή υπό κάποιο όνομα (σε πλήρη αντιστοιχία με το όνομα που χρησιμοποιείται από το μήνυμα ενδιαφέροντος για αναζήτησή της). Όταν ένα μήνυμα περιεχομένου φτάσει σε κάποιον κόμβο του δικτύου, ο κόμβος αυτός πρώτα διενεργεί έναν έλεγχο στην τοπική κρυφή μνήμη για να την ενημερώσει, τόσο εφόσον το μήνυμα δεν υπάρχει, όσο και αν υπάρχει, για τη νέα άφιξή του, ώστε να ενεργήσει σωστά και ο αλγόριθμος αντικατάστασης. Στη συνέχεια ελέγχεται ο πίνακας PIT για να διαπιστωθεί αν εκκρεμούν μηνύματα ενδιαφέροντος για το εν λόγω μήνυμα περιεχομένου. Υπό κανονικές συνθήκες, θα πρέπει να βρεθεί εγγραφή σε αυτόν τον πίνακα (αλλιώς δεν υπήρχε λόγος να φτάσει το μήνυμα περιεχομένου σε αυτόν τον κόμβο), η οποία θα περιλαμβάνει μια λίστα από faces από όπου θα πρέπει να προωθηθεί

το μήνυμα. Αν δε βρεθεί εγγραφεί στο PIT, το μήνυμα απορρίπτεται.

Από τις παραπάνω περιγραφές είναι προφανές ότι η διαδρομή που ακολουθεί ένα μήνυμα περιεχομένου είναι η ακριβώς αντίστροφη από αυτήν που ακολουθεί το αντίστοιχο μήνυμα ενδιαφέροντος, αφού το τελευταίο αφήνει ουσιαστικά το ίχνος του στον πίνακα PIT του κάθε κόμβου από τον οποίο περνάει και τον οποίο πίνακα συμβουλευεται ο εκάστοτε κόμβος για να χειριστεί τα μηνύματα περιεχομένου.

Η αρχιτεκτονική αυτή του CCNx, με τις βασικές δομές δεδομένων και τους τύπους μηνυμάτων της, επιλέχθηκε ως μοντέλο αναφοράς στην παρούσα διατριβή, καθώς είναι η πλέον διαδεδομένη στη βιβλιογραφία. Έτσι, ο προσομοιωτής JCCNxSim που περιγράφεται στην επόμενη ενότητα, υλοποιεί τις βασικές δομές δεδομένων και τύπους μηνυμάτων που παρουσιάστηκαν σε αυτήν την ενότητα, προκειμένου να γίνει η πειραματική εκτίμηση των επιδόσεων των επεκτάσεων που προτείνονται σε μια ρεαλιστική βάση.

2.4 Ο Προσομοιωτής JCCNxSim

Για τη διεξαγωγή των πειραμάτων που αφορούν τα πρωτόκολλα, τα οποία παρουσιάζονται στην παρούσα διατριβή στο πλαίσιο των πληροφοριοκεντρικών δικτύων, έγινε χρήση του προσομοιωτή JCCNxSim, ο οποίος αναπτύχθηκε ειδικά για αυτό το σκοπό. Η υλοποίησή του βασίστηκε στον προσομοιωτή PeerSim [21] [22], αρχικά σχεδιασμένο για υλοποίηση προσομοιώσεων για ομότιμα (peer-to-peer - P2P) δίκτυα.

2.4.1 Το Μεσιμικό Προσομοίωσης του PeerSim

Ο PeerSim είναι υλοποιημένος στη γλώσσα προγραμματισμού Java και παρέχεται ως μεσιμικό (middleware), προκειμένου να γίνει η ανάπτυξη των πρωτοκόλλων με βάση τους μηχανισμούς και το περιβάλλον εκτέλεσης που παρέχει. Υποστηρίζει δύο τύπους εκτελέσεων των πειραμάτων: κυκλικές (cycle driven) και οδηγούμενες από γεγονότα (event driven). Η κυκλική εκτέλεση αναφέρεται σε περιοδική εκτέλεση μιας διαδικασίας. Έτσι, σε κάθε κύκλο του πειράματος, ο προσομοιωτής κάνει κλήσεις των εν λόγω διαδικασιών, οι οποίες προγραμματίζονται από το χρήστη και υλοποιούν μια συγκεκριμένη διεπαφή παρεχόμενη από τον PeerSim, ώστε ο τελευταίος να είναι σε θέση να τις διαχωρίσει. Η εκτέλεση οδηγούμενη από γεγονότα από την άλλη αναφέρεται σε εκτέλεση διαδικασιών, οι οποίες ενεργοποιούνται όταν συμβεί κάποιο γεγονός, όπως η άφιξη ενός μηνύματος. Σε πλήρη αντιστοιχία με την προηγούμενη περίπτωση, ο PeerSim παρέχει μια διεπαφή, η οποία θα πρέπει να υλοποιηθεί από το χρήστη. Σε αυτήν την υλοποίηση θα πρέπει να περιγράφεται ο τρόπος χειρισμού των γεγονότων που λαμβάνουν χώρα, συνήθως προχωρώντας σε κλήσεις διαδικασιών που αφορούν το εκάστοτε πρωτόκολλο.

Επιπλέον των μηχανισμών εκτέλεσης των πρωτοκόλλων, ο PeerSim παρέχει και

διεπαφή που επιτρέπει στο χρήστη να ορίσει διαδικασίες αρχικοποίησης και περιοδικού ελέγχου των πειραμάτων, με τον τελευταίο να αφορά στη συνηθέστερη περίπτωση τη συλλογή ενδιάμεσων μετρήσεων. Ο χρήστης, υλοποιώντας την εν λόγω διεπαφή, μπορεί να ορίσει τις ενέργειες που θα πρέπει να γίνονται από το περιβάλλον εκτέλεσης της προσομοίωσης κατά το χρόνο κλήσης των αντίστοιχων μεθόδων.

Κάθε κόμβος του δικτύου που προσομοιώνεται αναπαρίσταται μέσω της διεπαφής Node, η οποία παρέχει τις βασικές μεθόδους προκειμένου να είναι σε θέση ο χρήστης να πραγματοποιήσει θεμελιώδεις πράξεις πάνω στους κόμβους. Τέτοιες πράξεις μεταξύ άλλων περιλαμβάνουν την πρόσβαση σε πρωτόκολλα που υλοποιούνται σε αυτόν τον κόμβο καθώς και την κατάστασή του (ενεργός/ανενεργός). Οι κόμβοι δημιουργούνται κατά την έναρξη της εκτέλεσης της προσομοίωσης και αναφορές σε αυτούς ανακτώνται μέσω της στατικής κλάσης Network.

Τα πρωτόκολλα ακολουθούν μια γενική υλοποίηση όπως περιγράφηκε νωρίτερα και δεν αφορούν τον κάθε κόμβο ξεχωριστά. Στην πραγματικότητα, κάθε κόμβος είναι δυνατό να φέρει την υλοποίηση ενός ή περισσότερων πρωτοκόλλων. Αυτό καθίσταται εφικτό μέσω της εσωτερικής αναπαράστασης κάθε πρωτοκόλλου με ένα μοναδικό αναγνωριστικό, το οποίο ανατίθεται από τον ίδιο τον προσομοιωτή, σε συνδυασμό και με το αρχείο διαχείρισης της προσομοίωσης, το οποίο περιγράφεται παρακάτω. Το μεσισμικό του προσομοιωτή δίνει τη δυνατότητα στο χρήστη να πάρει μια ανεξάρτητη οντότητα του εκάστοτε πρωτοκόλλου σε κάποιο κόμβο, μέσω των μεθόδων που παρέχει η διεπαφή Node.

Για την αποστολή μηνυμάτων μεταξύ των κόμβων, ο PeerSim παρέχει τη διεπαφή Transport, η οποία ορίζει δύο βασικές μεθόδους που αφορούν την αποστολή ενός μηνύματος αυτή καθαυτή καθώς και την καθυστέρηση μετάδοσης ενός μηνύματος ανάμεσα σε δύο κόμβους του δικτύου. Συγκεκριμένα, ως προς την αποστολή των μηνυμάτων, ο χρήστης ορίζει τον κόμβο πηγή και προορισμό, το μήνυμα και το πρωτόκολλο που θα πρέπει να χειριστεί το μήνυμα στον παραλήπτη. Αναφορικά με την καθυστέρηση μετάδοσης, η εν λόγω διεπαφή παρέχει μέθοδο που επιστρέφει το χρόνο σε χρονοθυρίδες (ο PeerSim μετρά το χρόνο με βάση χρονοθυρίδες) που απαιτείται για την αποστολή ενός μηνύματος ανάμεσα σε δύο κόμβους. Ο PeerSim παρέχει ενδεικτικές υλοποιήσεις της διεπαφής Transport, αλλά όπως με κάθε άλλο πρωτόκολλο, ο χρήστης έχει τη δυνατότητα να ορίσει τα δικά του. Στην παρούσα διατριβή ορίστηκαν αντίστοιχα πρωτόκολλα για υποστήριξη επικοινωνίας P2P και ροής πολυμεσικού περιεχομένου (multimedia content streaming) πάνω από πληροφοριοκεντρικά δίκτυα.

Για τη φόρτωση των παραμέτρων της προσομοίωσης στη μηχανή εκτέλεσης του PeerSim παρέχεται μια δομή αρχείου διαχείρισης, ως αρχείο απλού κειμένου, με βάση την οποία τα πρωτόκολλα/μηχανισμοί που ορίζονται από το χρήστη δύνανται να λαμβάνουν υπόψη τις εκάστοτε προτιμήσεις των χρηστών για την εκτέλεση της προσομοίωσης. Τέτοιες παράμετροι μπορεί να αφορούν τη χρήση συγκεκριμένων πρωτοκόλλων, χρονικές στιγμές εκτέλεσης και άλλες παραμέτρους που αφορούν την κάθε υλοποίηση.

2.4.2 Διάρθρωση και Υλοποίηση του JCCNxSim

Ο προσομοιωτής JCCNxSim υλοποιήθηκε πάνω από το μεσισμικό του PeerSim, εκμεταλλευόμενος τις δυνατότητες που παρέχει ο τελευταίος. Η επιλογή του PeerSim βασίστηκε στο γεγονός ότι τα πρωτόκολλα που χρησιμοποιήθηκαν για τις δοκιμές αφορούσαν αφενός επικοινωνία ομότιμων δικτύων, για τα οποία έχει άλλωστε σχεδιαστεί ο PeerSim, και αφετέρου ροή πολυμεσικού περιεχομένου, η οποία δύναται να προσομοιωθεί σε ποικίλα περιβάλλοντα.

Η υλοποίηση του JCCNxSim έγινε σε δύο επίπεδα, σε αντιστοιχία με το μεσισμικό του PeerSim. Το πρώτο επίπεδο αφορά το δίκτυο, τους τύπους των μηνυμάτων και τις συσκευές του δικτύου, με όλες τις δομές δεδομένων που αυτές υλοποιούν. Το δεύτερο επίπεδο αφορά τις επεκτάσεις που χρειάστηκε να γίνουν προκειμένου να υποστηριχθούν οι προσομοιώσεις για τα προτεινόμενα πρωτόκολλα που αφορούν ομότιμα δίκτυα και ροή πολυμεσικού περιεχομένου. Οι επεκτάσεις αυτές αφορούσαν τόσο τύπους μηνυμάτων, όσο και νέες υλοποιήσεις πρωτοκόλλων μεταφοράς (σε επίπεδο μεσισμικού του PeerSim και όχι δικτύου) προκειμένου να υποστηριχθούν διαφανώς οι νέες διαδικασίες.

2.4.2.1 Υλοποίηση Βασικών Δομών και Μηχανισμών του CCNx

Η προσομοίωση του CCNx απαιτεί σε πρώτο στάδιο την εξομοίωση των βασικών δομών δεδομένων και μηνυμάτων που χρησιμοποιεί. Έτσι, για τον JCCNxSim υλοποιήθηκαν δύο κλάσεις, για το Interest και το Content Object αντίστοιχα, οι οποίες δίνουν πρόσβαση στα δεδομένα αυτών των μηνυμάτων. Στον προσομοιωτή δε χρησιμοποιήθηκε όλη η πληροφορία που μπορεί να φέρουν αυτά τα μηνύματα με βάση τις προδιαγραφές, αλλά περιορίζεται σε εκείνη που απαιτείται για τις προσομοιώσεις που πραγματοποιήθηκαν, ήτοι το όνομα του πακέτου που μεταφέρεται καθώς και το περιεχόμενο αυτό καθαυτό για την περίπτωση του Content Object. Για λόγους παρακολούθησης της εξέλιξης του πειράματος, τα μηνύματα μεταφέρουν επίσης την ακόλουθη πληροφορία: κόμβος πηγή, κόμβος προορισμός, αριθμός κόμβων που έχει επισκεφθεί το μήνυμα και μια ένδειξη για το αν το μήνυμα έχει εξυπηρετηθεί από την κρυφή μνήμη.

Επιπλέον των μηνυμάτων, ο JCCNxSim υλοποιεί και τις βασικές δομές ενός CCNx κόμβου: την αποθήκη περιεχομένου (CS), τον πίνακα εκκρεμών ενδιαφερόντων (PIT) καθώς και τον πίνακα προώθησης περιεχομένου (FIB). Η αποθήκη περιεχομένου έχει υλοποιηθεί ως πίνακας κατακερματισμού πεπερασμένου μεγέθους. Υποστηρίζονται οι αλγόριθμοι αντικατάστασης του παλαιότερα χρησιμοποιημένου στοιχείου (Least Recently Used - LRU) και του αραιότερα χρησιμοποιημένου στοιχείου (Least Frequently Used - LFU). Το μέγεθος της αποθήκης περιεχομένου δίνεται δυναμικά για κάθε πείραμα μέσω του αρχείου διαχείρισης.

Για τη δομή του PIT, η προφανής επιλογή θα ήταν ένας πίνακας κατακερματισμού

ανάμεσα στο όνομα του περιεχομένου που αιτείται ο χρήστης και τη διεπαφή δικτύου (face) από την οποία προήλθε το σχετικό Interest μήνυμα. Αντί αυτού όμως, προτιμήθηκε να δημιουργηθεί μια λίστα για κάθε διεπαφή περιλαμβάνοντας τα εν λόγω ονόματα. Ο λόγος που έγινε αυτή η επιλογή ήταν η ευκολότερη διαχείριση της πληροφορίας σε επίπεδο προσομοιωτή. Έτσι, για την περίπτωση που φτάνουν σε ένα κόμβο του δικτύου μηνύματα Interest για το ίδιο περιεχόμενο (άρα με το ίδιο όνομα) αλλά προέρχονται από διαφορετικές διεπαφές δικτύου, θα απαιτούνταν πιο πολύπλοκη δομή από έναν απλό πίνακα κατακερματισμού με κλειδί το όνομα του περιεχομένου, δεδομένου ότι κάθε όνομα μπορεί να αντιστοιχεί σε πολλές διεπαφές δικτύου.

Η δομή του FIB που αποτελεί ουσιαστικά τον πίνακα δρομολόγησης του κάθε κόμβου υλοποιήθηκε με έναν πίνακα κατακερματισμού. Στην τρέχουσα υλοποίηση, υπάρχει αντιστοίχιση ολόκληρων των ονομάτων των στοιχείων πληροφορίας που δημοσιεύονται με τις διεπαφές δικτύου από τις οποίες πρέπει να προωθηθούν τα μηνύματα Interest προκειμένου να φτάσουν στο αντίστοιχο περιεχόμενο. Για την καλύτερη κλιμάκωση του πίνακα κατακερματισμού του FIB, το CCNx προτείνει τη χρήση προθεμάτων όπου θα γίνεται ταίριασμα μέγιστου προθέματος, ώστε να μη χρειάζεται να αποθηκεύονται όλα τα πιθανά ονόματα. Εν τούτοις, οι ανάγκες του προσομοιωτή για τα πλαίσια της τρέχουσας διατριβής δεν υπαγόρευαν μια τέτοια προσέγγιση και έτσι αφήνεται στο άμεσο μελλοντικό έργο. Για την ενημέρωση των πινάκων FIB, ορίστηκε ένας νέος τύπος μηνυμάτων (Content Advertisement) στα πλαίσια του προσομοιωτή. Αυτά τα μηνύματα στέλνονται από τους κόμβους των χρηστών, όταν οι τελευταίοι θέλουν να δημοσιεύσουν περιεχόμενο στο δίκτυο και μεταφέρουν πληροφορία για το όνομα που αντιστοιχεί στο δημοσιευόμενο περιεχόμενο καθώς και το πλήθος των δρομολογητών που έχει επισκεφθεί μέχρι να φτάσει στον εκάστοτε δρομολογητή (μήκος μονοπατιού).

Η λειτουργικότητα ενός CCNx κόμβου προσομοιώνεται από ειδικό πρωτόκολλο που αναπτύχθηκε συγκεκριμένα για αυτό το σκοπό. Λαμβάνοντας υπόψη του τις προδιαγραφές του CCNx, το εν λόγω πρωτόκολλο αναλαμβάνει το χειρισμό των τριών προαναφερθέντων τύπων μηνυμάτων (Interest, Content Object, Content Advertisement). Το πρωτόκολλο υλοποιήθηκε ως διαδικασία οδηγούμενη από γεγονότα (event-driven) με κάθε γεγονός να αντιστοιχεί στην άφιξη ενός από τα παραπάνω μηνύματα σε κάποιο κόμβο του δικτύου.

Συγκεκριμένα, η άφιξη ενός μηνύματος Content Advertisement ενεργοποιεί το μηχανισμό ενημέρωσης των πινάκων δρομολόγησης. Το πρωτόκολλο ορίζει ότι όταν ένας κόμβος δικτύου λαμβάνει ένα τέτοιο μήνυμα, ελέγχει αν υπάρχει ήδη εγγραφή στο FIB για αυτό το όνομα και αν δεν υπάρχει, τότε δημιουργεί την αντιστοιχία ανάμεσα στο όνομα και στη διεπαφή δικτύου από όπου έφτασε. Παράλληλα κρατά σε μια ξεχωριστή δομή το μήκος του μονοπατιού που ακολουθήθηκε μέχρι να φτάσει εκεί το μήνυμα. Έτσι, αν στο μέλλον φτάσει ένα επόμενο τέτοιο μήνυμα, επειδή ακολούθησε διαφορετική διαδρομή, θα επιλέξει το συντομότερο μονοπάτι. Στις προσομοιώσεις που πραγματοποιήθηκαν, η ενημέρωση των πινάκων δρομολόγησης γινόταν κατά τη διάρκεια της αρχικοποίησης του προ-

σομοιωτή, χωρίς όμως αυτό να είναι απαίτηση του λογισμικού, το οποίο υποστηρίζει την ανανέωση των πινάκων δρομολόγησης και σε πραγματικό χρόνο.

Για τα μηνύματα Content Object, η βασική έκδοση του προσομοιωτή λαμβάνει υπόψη δύο πιθανές περιπτώσεις: την ύπαρξη ή μη της αποθήκης περιεχομένου CS. Στην πρώτη περίπτωση, το μήνυμα αποθηκεύεται αρχικά στην αποθήκη περιεχομένου, με βάση τις μεθόδους που παρέχονται για την υλοποίηση και του αντίστοιχου αλγορίθμου αντικατάστασης, όπως εξηγήθηκε νωρίτερα σε αυτήν την Ενότητα. Στη συνέχεια, όπως και στη δεύτερη περίπτωση της μη ύπαρξης CS, γίνεται έλεγχος στον πίνακα PIT, όπως αυτός υλοποιείται στα πλαίσια της κάθε διεπαφής δικτύου ξεχωριστά, και προωθείται μέσω των διεπαφών δικτύου που έχουν εκκρεμή μηνύματα Interest για περιεχόμενο με το ίδιο όνομα.

Ακολουθώντας τις προδιαγραφές του CCNx, όταν παραληφθεί ένα μήνυμα Interest, πρώτα ελέγχεται η αποθήκη περιεχομένου για εγγραφή με το ίδιο όνομα, εφόσον πρόκειται για προσομοίωση που υποστηρίζει την ύπαρξη CS. Αν δε βρεθεί τέτοια εγγραφή, τότε το πρωτόκολλο συμβουλευτεί τον πίνακα FIB, ώστε να προωθήσει το Interest στον κατάλληλο κόμβο του δικτύου. Παράλληλα, στην τελευταία περίπτωση, ενημερώνεται και η δομή PIT της αντίστοιχης διεπαφής δικτύου.

2.4.2.2 Υλοποίηση Επεκτάσεων του CCNx

Οι επεκτάσεις που υλοποιήθηκαν στο CCNx αφορούν υλοποιήσεις για την υποστήριξη προσομοιώσεων για ομότιμα δίκτυα και ροή πολυμεσικού περιεχομένου. Οι επεκτάσεις αυτές αφορούν τόσο την υλοποίηση του πρωτοκόλλου CCNx που περιγράφηκε στην προηγούμενη ενότητα, ώστε το δίκτυο να είναι σε θέση να χειρίζεται τις επεκτάσεις των μηνυμάτων Interest που θα περιγραφούν σε επόμενες ενότητες αφιερωμένες στην περιγραφή των εν λόγω μηχανισμών όσο και πρωτόκολλα Transport (σε επίπεδο προσομοιωτή και όχι στοίβας δικτυακών πρωτοκόλλων).

Συγκεκριμένα, για την υποστήριξη της αποδοτικής διακίνησης περιεχομένου με χρήση ομότιμων πρωτοκόλλων, τα οποία εκμεταλλεύονται και τις εγγενείς δυνατότητες του δικτύου για προσωρινή αποθήκευση του περιεχομένου, εισήχθη μια επέκταση του μηνύματος Interest, το P2PInterest. Ο χειρισμός αυτού του μηνύματος από το δίκτυο περιγράφεται με λεπτομέρεια στην Ενότητα 2.5, όπου αναλύεται ο εν λόγω μηχανισμός. Στο επίπεδο της υλοποίησης, ο νέος τύπος μηνύματος σχηματίζεται από ένα νέο πρωτόκολλο σε επίπεδο Transport, το οποίο καλείται από τις αντίστοιχες μεθόδους του επιπέδου εφαρμογής. Το επίπεδο του δικτύου αντιλαμβάνεται το P2PInterest ως ένα ξεχωριστό γεγονός και, ως εκ τούτου, το χειρίζεται ξεχωριστά.

Για τη δημιουργία ενός μηνύματος P2PInterest, απαιτείται να δοθούν οι κόμβοι πηγής και προορισμού, το όνομα του περιεχομένου που φέρει το εν λόγω μήνυμα, το μήνυμα αυτό καθαυτό και τέλος, το όνομα της υπηρεσίας που παρέχει ο κόμβος προορισμού προκειμένου να λαμβάνει P2P μηνύματα. Δεδομένου ότι πρόκειται για πρωτόκολλο

ομότιμης επικοινωνίας, η κάθε συμμετέχουσα συσκευή θεωρείται ότι δημοσιεύει στο δίκτυο το όνομα της υπηρεσίας, στην οποία θα πρέπει να δρομολογούνται τα πακέτα. Στην επέκταση του προσομοιωτή, αυτό είναι και το όνομα που χρησιμοποιείται για τη δρομολόγηση των εν λόγω μηνυμάτων, σε αντίθεση με ό,τι συμβαίνει με το τυπικό μοντέλο επικοινωνίας πελάτη-εξυπηρετητή, όπου χρησιμοποιείται το όνομα του περιεχομένου, όπως περιγράφηκε παραπάνω.

Σε πλήρη αντιστοιχία με την προηγούμενη περίπτωση των ομότιμων δικτύων, έτσι και για την υποστήριξη των επεκτάσεων αποτελεσματικότερης υποστήριξης ροής πολυμεσικού περιεχομένου, υλοποιήθηκε νέο πρωτόκολλο Transport, το οποίο παρέχει τη δυνατότητα δημιουργίας μηνυμάτων MultimediaInterest, όπως αυτά περιγράφονται στη σχετική ενότητα 2.6. Η επιπλέον υποστήριξη, σε σχέση με το βασικό πρωτόκολλο Transport που υλοποιήθηκε στον προσομοιωτή για την υποστήριξη του πρωτοκόλλου CCNx, είναι ότι περιλαμβάνει μια επιπλέον λογική παράμετρο, ώστε η εφαρμογή αναπαραγωγής του πολυμεσικού περιεχομένου να μπορεί να ενεργοποιεί την προτεινόμενη λειτουργικότητα.

2.5 Υποστήριξη Διαμοιρασμού Περιεχομένου σε Δίκτυα Ομότιμων Κόμβων

Το θεμελιώδες χαρακτηριστικό των πληροφοριοκεντρικών δικτύων, πάνω στο οποίο έχει εν πολλοίς βασιστεί το κίνητρο για τη χρήση τους, είναι η κρυφή μνήμη στους κόμβους του δικτύου, ώστε οι αιτήσεις να μην εξυπηρετούνται πάντα από τον κόμβο που παρέχει το περιεχόμενο, αλλά και από ενδιάμεσους κόμβους του δικτύου. Αυτή η εγγενής συμπεριφορά των πληροφοριοκεντρικών δικτύων θυμίζει δίκτυα ομότιμων κόμβων, όπου το εκατόστοτε πρωτόκολλο αποφασίζει από ποιόν κόμβο τελικά θα παραληφθεί το περιεχόμενο.

Ωστόσο, το παράδοξο είναι ότι τα ομότιμα πρωτόκολλα δεν μπορούν να αξιοποιήσουν την κρυφή μνήμη του δικτύου. Ο λόγος είναι ότι βασίζονται σε υπηρεσίες που παρέχουν οι συμμετέχοντες κόμβοι, ενώ είναι ευθύνη του πρωτοκόλλου να αποφασίσει από ποιον κόμβο θα ζητηθεί το περιεχόμενο και όχι του δικτύου. Έτσι, σε αυτήν την περίπτωση, το δίκτυο περιορίζεται στο να δρομολογεί απλά τα πακέτα προς μια υπηρεσία που υποδεικνύεται από τα ανώτερα στρώματα και, κατά συνέπεια, τα δεδομένα που διακινούνται θεωρούνται *ανώνυμα*, τουλάχιστον όπως υποστηρίζονται μέχρι στιγμής στη βιβλιογραφία [20] [23].

Η κίνηση για διαμοιρασμό αρχείων με BitTorrent για το 2013 προβλέπεται να αποτελέσει το 9,23% της συνολικής κίνησης στη Βόρεια Αμερική, το 17,36% στην Ευρώπη και το 21,66% στην Ασία-Ειρηνικό [24]. Συνεπώς, είναι προφανές ότι ο σχεδιασμός ενός δικτύου νέας γενιάς θα πρέπει να λαμβάνει σοβαρά υπόψη τις απαιτήσεις που ανακύπτουν από τα εν λόγω πρωτόκολλα.

Σε αυτήν την ενότητα, παρουσιάζεται μια επέκταση του πληροφοριοκεντρικού μο-

ντέλου δικτύου του CCNx, ώστε να διαχειρίζεται με πιο αποδοτικό τρόπο πληροφορία, η οποία διακινείται με το πρωτόκολλο BitTorrent [25]. Επιπλέον, δίνονται πειραματικά αποτελέσματα που προέκυψαν από προσομοιώσεις, οι οποίες πραγματοποιήθηκαν με τον προσομοιωτή JCCNxSim και αναδεικνύουν την καλύτερη χρήση της κρυφής μνήμης μεταξύ μιας τυπικής αρχιτεκτονικής πελάτη-εξυπηρετητή και ομότιμου δικτύου.

2.5.1 Το Πρωτόκολλο Διαμοιρασμού Περιεχομένου BitTorrent

Το BitTorrent αποτελεί το πιο διαδεδομένο πρωτόκολλο διαμοιρασμού περιεχομένου και για το λόγο αυτό θα χρησιμοποιηθεί σε αυτήν την ενότητα, ως μοντέλο αναφοράς. Η βασική αρχή που διέπει τη λειτουργία του και αποτελεί και τον κύριο λόγο της επιτυχίας του είναι ότι κάθε κόμβος που κατεβάζει κομμάτια του περιεχομένου υποχρεούται να παρέχει στους υπόλοιπους συμμετέχοντες αυτά που έχει ήδη κατεβάσει. Περιλαμβάνει τρεις τύπους κόμβων ανάλογα με τη θέση τους στην όλη αρχιτεκτονική και την ποσότητα του περιεχομένου που διαθέτουν προς διαμοιρασμό. Ο πρώτος τύπος είναι οι κόμβοι που έχουν ολοκληρω το περιεχόμενο και το μοιράζουν στους υπόλοιπους (*seeders*). Ο δεύτερος είναι οι κόμβοι, οι οποίοι κατεβάζουν και, ταυτόχρονα, διαμοιράζουν το περιεχόμενο (*leechers*)¹. Ο τρίτος τύπος είναι ο κόμβος που διαθέτει πληροφορίες για τους *seeders* και τους *leechers*, με τους τελευταίους να επικοινωνούν περιοδικά με αυτόν προκειμένου να κατεβάσουν μια λίστα με τους άλλους κόμβους από τους οποίους μπορούν να κατεβάσουν κομμάτια του περιεχομένου (*tracker*). Ο *tracker* μπορεί να είναι ένας μοναδικός κόμβος ή να είναι και μια κατακευματισμένη υπηρεσία (*trackerless torrents*).

Το περιεχόμενο που διαμοιράζεται στο BitTorrent χωρίζεται σε κομμάτια (*pieces*) και κάθε κομμάτι σε επιμέρους τμήματα (*blocks*). Αν και δεν υπάρχει κάποιος συγκεκριμένος, αυστηρός κανόνας, γενικά το μέγεθος ενός κομματιού ποικίλει από 16KB μέχρι 128KB, ανάλογα και με τη σύνδεση που προτείνεται να υποστηρίξει το διαμοιρασμό. Ο δε χωρισμός σε επιμέρους τμήματα είναι εικονικός και αφορά τις αιτήσεις που πραγματοποιεί ο εκάστοτε χρήστης, όπου δίνεται ρητά το μέγεθος του περιεχομένου που αιτείται.

Για να συμμετέχει κάποιος χρήστης στο διαμοιρασμό ενός αρχείου (ή ομάδας αρχείων) με το BitTorrent, πρέπει αρχικά να διαθέτει το αντίστοιχο αρχείο *torrent*. Αυτό περιλαμβάνει πληροφορίες όπως η διεύθυνση (URL) του *tracker*, ο αριθμός κομματιών που είναι χωρισμένο το περιεχόμενο καθώς και το μέγεθος του καθενός τέτοιου κομματιού. Αφού λάβει αυτό το αρχείο, ο χρήστης θα επικοινωνήσει με τον *tracker* προκειμένου να λάβει τη λίστα με τους συμμετέχοντες κόμβους και να ξεκινήσει να κατεβάζει και να μοιράζει περιεχόμενο.

Το BitTorrent ορίζει διάφορους τύπους μηνυμάτων [26], ωστόσο στην παρούσα ενότητα παρουσιάζονται μόνο εκείνοι που έχουν χρειαστεί στις προσομοιώσεις που υλοποιούν

¹Αν και γενικά ο όρος *leecher* αφορά σε κόμβους που εμφανίζουν εγωιστική συμπεριφορά (ήτοι μόνο κατεβάζουν), το BitTorrent θεωρεί όλους τους κόμβους που δεν έχουν ολοκληρω το περιεχόμενο ως *leechers*.

το πρωτόκολλο. Για να πραγματοποιήσει μια αίτηση, ο χρήστης στέλνει ένα μήνυμα request σε κάποιον άλλο κόμβο, στο οποίο αναφέρει το κομμάτι (piece) που επιθυμεί να λάβει, καθώς και το επιμέρους τμήμα του. Η απάντηση σε μια αίτηση ενθυλακώνεται σε ένα μήνυμα piece, το οποίο περιλαμβάνει τον αριθμό του κομματιού καθώς και το επιμέρους τμήμα του που επισυνάπτεται στο μήνυμα. Όταν έχει ολοκληρώσει το κατέβασμα ενός κομματιού, ο κόμβος στέλνει στους υπόλοιπους ένα μήνυμα have με τον αριθμό αυτού του κομματιού. Κάθε κόμβος που λαμβάνει αυτό το μήνυμα και δεν έχει ή δεν κατεβάζει ήδη αυτό το κομμάτι από κάποιον άλλο, το αιτείται από τον αποστολέα του have. Στην τρέχουσα υλοποίηση στον JCCNxSim, κάθε κόμβος κατεβάζει ολόκληρο το κομμάτι από αυτόν τον κόμβο που ξεκίνησε, ενώ όλοι ξεκινούν από κάποιον κόμβο seeder. Επιπλέον, η σειρά με την οποία επιλέγεται το κομμάτι που αιτείται κάθε φορά ένας κόμβος είναι τυχαία, εκτός αν η αίτηση είναι αποτέλεσμα παραλαβής μηνύματος have.

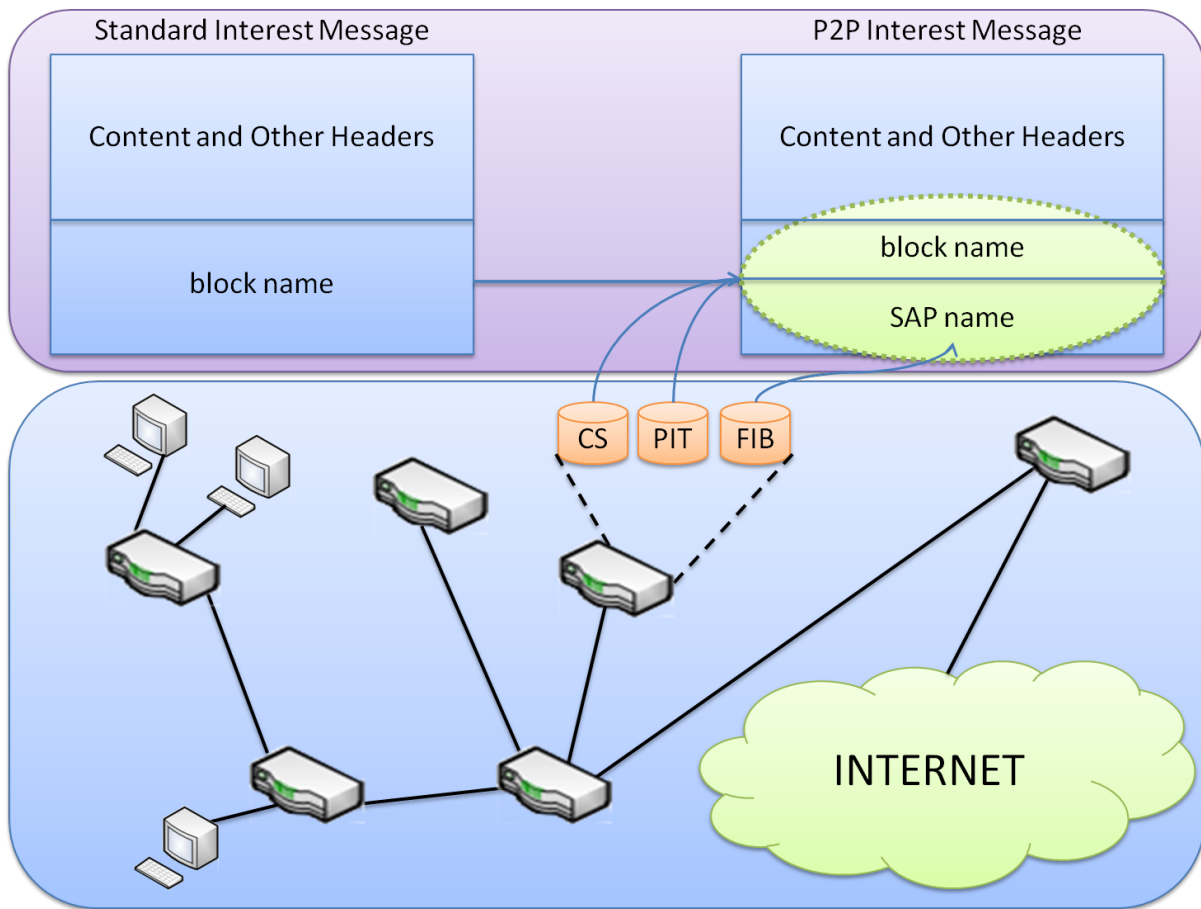
2.5.2 Αρχιτεκτονική Ομότιμης Επικοινωνίας σε Πληροφοριοκεντρικό Δίκτυο

Η επέκταση του πληροφοριοκεντρικού δικτύου CCNx για την υποστήριξη ομότιμης επικοινωνίας έχει ως στόχο τη διατήρηση των ιδιαίτερων χαρακτηριστικών των ομότιμων πρωτοκόλλων, ενώ ταυτόχρονα γίνεται εκμετάλλευση των δυνατοτήτων που παρέχει ένα πληροφοριοκεντρικό δίκτυο και δει της κρυφής μνήμης. Η βασική ιδέα που διέπει αυτήν την επέκταση είναι ότι κάθε πακέτο Interest φέρει δύο ονόματα: το όνομα του περιεχομένου καθώς και το όνομα της υπηρεσίας του κόμβου από τον οποίο θα πρέπει να εξυπηρετηθεί.

Στην πραγματικότητα, όμως, το όνομα της υπηρεσίας δεν είναι παρά το όνομα στο οποίο θα πρέπει να βασιστεί ο αλγόριθμος δρομολόγησης προκειμένου να στείλει την αίτηση στο σωστό κόμβο εξυπηρέτησης περιεχομένου. Άλλωστε, με μια αφαιρετική ματιά, σε ένα ομότιμο δίκτυο, όλοι οι συμμετέχοντες κόμβοι είναι εν δυνάμει κόμβοι εξυπηρέτησης του περιεχομένου που έχουν κατεβάσει.

Λαμβάνοντας υπόψη το γεγονός αυτό, μια λύση που χρίζει περαιτέρω διερεύνησης είναι κάθε κόμβος που κατεβάζει ένα κομμάτι του περιεχομένου, ταυτόχρονα να το δημοσιεύει στο δίκτυο, ώστε το τελευταίο να αφομοιώσει χαρακτηριστικά ομότιμης επικοινωνίας. Ωστόσο, για να γίνει κάτι τέτοιο, θα πρέπει να χρησιμοποιηθεί διαφορετικό όνομα σε κάθε κόμβο, καθώς διαφορετικά θα υπήρχε πρόβλημα δρομολόγησης. Επιπλέον, όσο μεγαλώνει ο αριθμός των κόμβων που διαμοιράζονται το περιεχόμενο, θα γίνεται και μεγαλύτερος ο αριθμός αυτών των ονομάτων, πράγμα που θα μπορούσε να οδηγήσει σε προβλήματα κλιμάκωσης των πινάκων δρομολόγησης, τα οποία μάλιστα υφίστανται έτσι κι αλλιώς. Τέλος, ακόμα και αν βρισκόταν λύση για τα υπόλοιπα, θα έπρεπε το δίκτυο να αποφασίζει από μόνο του πού θα δρομολογήσει κάθε αίτηση, καταργώντας έτσι στην πράξη την ύπαρξη των υπερκείμενων δικτύων που δημιουργούν τα πρωτόκολλα ομότιμης επικοινωνίας προκειμένου να λειτουργήσουν με αποδοτικό τρόπο.

Το Σχήμα 2 δείχνει το τυπικό μήνυμα Interest καθώς και την επέκτασή του για την



Σχήμα 2: Υποστήριξη Ομότιμης Επικοινωνίας στο Δίκτυο

υποστήριξη της ομότιμης επικοινωνίας. Για να υποστηριχθεί η δρομολόγηση με βάση το όνομα της υπηρεσίας (Service Access Point - SAP) του κόμβου επιλογής του πρωτοκόλλου ομότιμης επικοινωνίας, θα πρέπει αυτό να περιλαμβάνεται μήνυμα Interest, ώστε να αξιοποιηθεί από τη δομή FIB που είναι υπεύθυνη για την προώθηση των πακέτων. Από την άλλη, για να γίνει εξυπηρέτηση από την κρυφή μνήμη καθώς και έλεγχος για το ποια μηνύματα Interest έχουν ήδη προωθηθεί (ώστε να μη γίνει επαναποστολή τους) σε έναν κόμβο του δικτύου, χρειάζεται το όνομα του επιμέρους τμήματος του περιεχομένου που περιλαμβάνει η αίτηση. Έτσι και αυτό, συμπεριλαμβάνεται στο μήνυμα Interest.

Για να μπορέσει να λειτουργήσει το προτεινόμενο σχήμα, υπάρχουν δύο βασικές προϋποθέσεις. Η πρώτη αφορά το σχήμα ονομάτων που θα χρησιμοποιηθεί για τα επιμέρους τμήματα περιεχομένου (blocks), προκειμένου να υπάρχει συνέπεια μεταξύ των χρηστών που συμμετέχουν στο διαμοιρασμό τους. Δεδομένου ότι δεν πρόκειται για περιεχόμενο, το οποίο δημοσιεύεται όπως είναι, το σχήμα ονομάτων δεν είναι απαραίτητο να ακολουθεί αυστηρά τον ορισμό του δικτύου· εντούτοις, είναι προφανές ότι θα πρέπει να είναι μοναδικό για κάθε torrent. Η δεύτερη προϋπόθεση έχει να κάνει με το μέγεθος των επιμέρους τμημάτων, τα οποία αποθηκεύονται στην κρυφή μνήμη. Συγκεκριμένα, θα πρέπει

να υπάρχει μια καθολική προσέγγιση σε κάθε torrent που διαμοιράζεται και να μη ζητά ο κάθε χρήστης ό,τι μέγεθος θέλει, καθώς έτσι δε θα μπορούσε να λειτουργήσει αποδοτικά η κρυφή μνήμη. Αν και θα μπορούσε να επιστρέφει ό,τι έχει αποθηκευμένο, ανεξαρτήτως της αίτησης, κάτι τέτοιο θα εισήγαγε επιπλέον πολυπλοκότητα και για αυτό το λόγο προτείνεται ο αυστηρός εξαρχής καθορισμός του μεγέθους στο αρχείο torrent που κατεβάζει αρχικά ο κάθε χρήστης, ώστε όλοι να αιτούνται τμημάτων ίδιου μεγέθους.

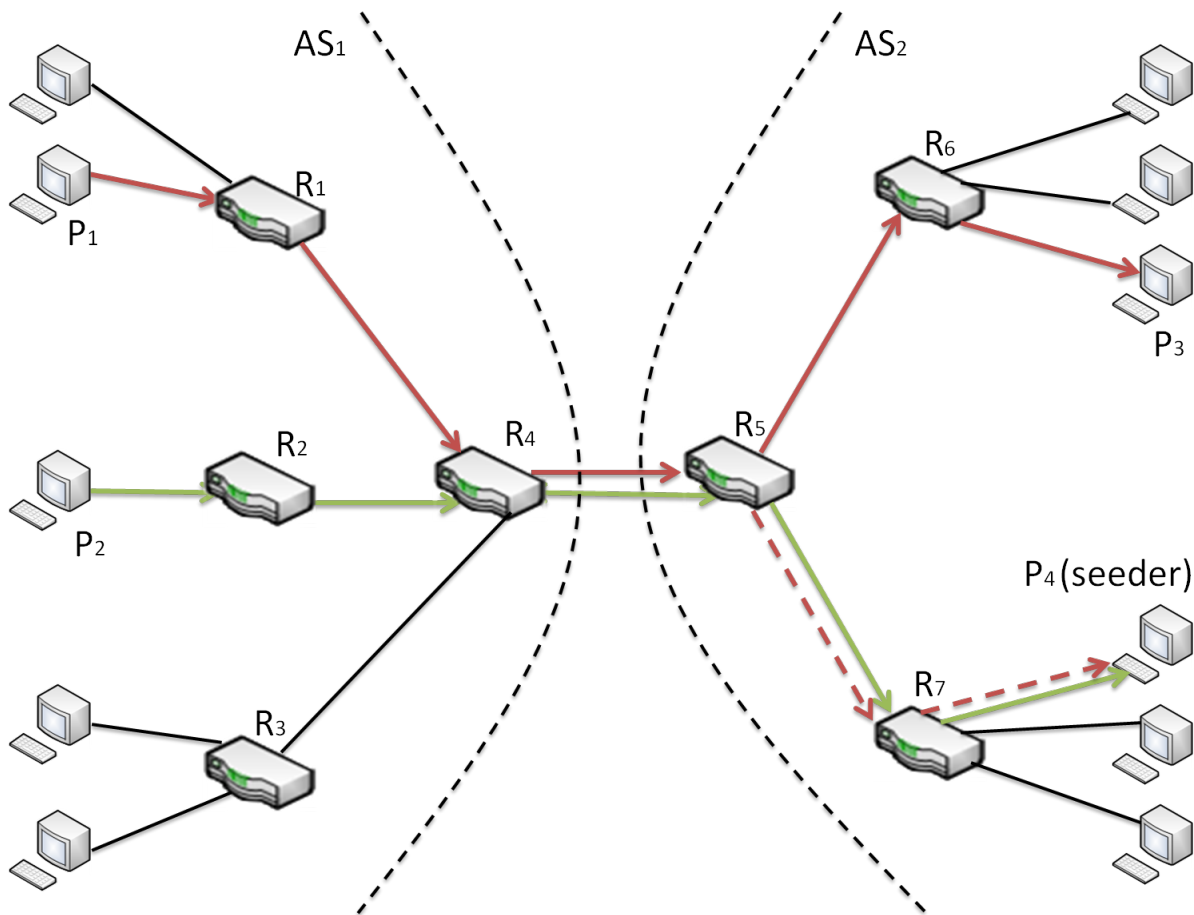
Όταν κάποιος δρομολογητής λάβει μήνυμα Interest, τότε αρχικά εξάγει το όνομα του τμήματος που αιτείται ο χρήστης και ελέγχει, αρχικά την κρυφή μνήμη και, αν δε βρει κάτι εκεί, τον πίνακα εκκρεμών ενδιαφερόντων (PIT). Ακολουθεί, δηλαδή, την τυπική διαδικασία αίτησης περιεχομένου από κάποιον κόμβο εξυπηρέτησης. Σε περίπτωση που πρέπει να προωθήσει το μήνυμα προς αυτόν τον κόμβο, εξάγει το όνομα της υπηρεσίας του από το Interest και, με βάση την πληροφορία του πίνακα δρομολόγησης FIB, το προωθεί στον κατάλληλο κόμβο. Με αυτόν τον τρόπο, εκμεταλλεύεται αφενός τα ιδιαίτερα χαρακτηριστικά του πληροφοριοκεντρικού δικτύου, ενώ αφετέρου ακολουθεί τη δρομολόγηση υπερκείμενου δικτύου που προδιαγράφει το ανώτερο πρωτόκολλο ομότιμης επικοινωνίας.

Το πλεονέκτημα αυτής της προσέγγισης, σε σύγκριση με μια τυπική αρχιτεκτονική πελάτη-εξυπηρετητή είναι ότι, λόγω της κατανομής του φόρτου σε διάφορους κόμβους σε επίπεδο υπερκείμενου δικτύου, ομοίως κατανέμονται σε περισσότερους δρομολογητές τα τμήματα του περιεχομένου που κατεβάζουν οι χρήστες, λόγω των διαφορετικών διαδρομών που ακολουθούνται. Όπως φαίνεται και στην Ενότητα 2.5.3, το γεγονός αυτό έχει θετική επίδραση στην πιθανότητα εύρεσης περιεχομένου στην κρυφή μνήμη, η οποία είναι ιδιαίτερα αυξημένη σε σχέση με το μοντέλο επικοινωνίας πελάτη-εξυπηρετητή.

Ο παραπάνω ισχυρισμός παρουσιάζεται μέσω ενός παραδείγματος που δίνεται στο Σχήμα 3, το οποίο παρουσιάζει μια εικόνα ενός δικτύου με δύο αυτόνομα συστήματα, των οποίων κάποιοι χρήστες επικοινωνούν με πρωτόκολλο ομότιμης επικοινωνίας. Έστω ότι το εν λόγω πρωτόκολλο αποφασίζει ότι ο κόμβος P_1 πρέπει να ζητήσει το περιεχόμενο από τον κόμβο P_3 και ο P_2 από τον P_4 , ο οποίος είναι και seeder και άρα ισοδύναμος με έναν κόμβο εξυπηρέτησης περιεχομένου σε αρχιτεκτονική πελάτη-εξυπηρετητή². Έστω, επίσης, ότι ανάμεσα στους δρομολογητές R_4 και R_5 παρεμβάλλεται κάποιο δημόσιο δίκτυο, όπου ο κάθε πάροχος πληρώνει με βάση τη χρήση του (όπως γίνεται σήμερα με το διαδίκτυο). Τα βέλη του σχήματος υποδεικνύουν τη διαδρομή που θα ακολουθήσουν οι αιτήσεις, ενώ οι απαντήσεις, ακολουθώντας τις προδιαγραφές του CCNx, θα ακολουθήσουν την ακριβώς αντίθετη τροχιά.

Αν ακολουθούσαν μια αρχιτεκτονική πελάτη-εξυπηρετητή, τότε και η αίτηση του P_1 θα δρομολογούνταν προς τον P_4 που είναι ο seeder, όπως υποδεικνύει και η κόκκινη διακεκομμένη γραμμή. Σε μια τέτοια περίπτωση, το περιεχόμενο θα αποθηκευόταν προσωρινά στην κρυφή μνήμη των δρομολογητών $\{R_1, R_4, R_5, R_7\}$. Αντίστοιχα, για την επικοι-

²Η ισοδυναμία αυτή ισχύει φυσικά σε περίπτωση που υπάρχει ένας μόνο seeder, αλλά η εν λόγω θεώρηση βοηθά τη σύγκριση των δύο μοντέλων επικοινωνίας χωρίς βλάβη της γενικότητας.



Σχήμα 3: Παράδειγμα Αίτησης σε Ομότιμη Επικοινωνία

ωνία του P_2 με τον P_4 οι δρομολογητές που διατηρούν το περιεχόμενο στην κρυφή τους μνήμη είναι οι $\{R_2, R_4, R_5, R_7\}$. Συνεπώς, με το μοντέλο πελάτη-εξυπηρετητή, το περιεχόμενο θα βρίσκεται στην κρυφή μνήμη των δρομολογητών που ανήκουν στην ένωση των δύο παραπάνω συνόλων, ήτοι $\{R_1, R_2, R_4, R_5, R_7\}$. Ωστόσο, στην περίπτωση της ομότιμης επικοινωνίας, όπου ο P_1 παίρνει το περιεχόμενο από τον P_3 , οι δρομολογητές για αυτήν τη διαδρομή είναι οι $\{R_1, R_4, R_5, R_6\}$. Άρα, τελικά, το περιεχόμενο θα περιλαμβάνεται στις κρυφές μνήμες των $\{R_1, R_2, R_4, R_5, R_6, R_7\}$, δηλαδή θα υπάρχει ένας επιπλέον δρομολογητής (R_6) που θα έχει το εν λόγω τμήμα του περιεχομένου στην κρυφή του μνήμη.

2.5.3 Πειραματική Εκτίμηση Επιδόσεων της Προτεινόμενης Αρχιτεκτονικής

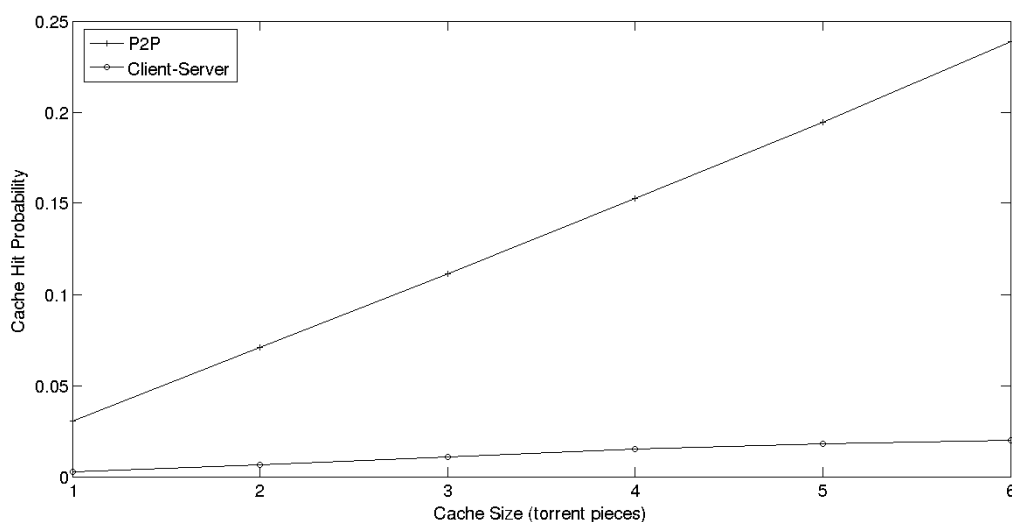
Η εκτίμηση της απόδοσης της προτεινόμενης αρχιτεκτονικής για υποστήριξη πρωτοκόλλων ομότιμης επικοινωνίας πάνω από πληροφοριοκεντρικό δίκτυο έγινε με βάση πειραματικά αποτελέσματα που προέκυψαν χρησιμοποιώντας τον προσομοιωτή JCCNxSim με τις αντίστοιχες επεκτάσεις του. Στο επίπεδο εφαρμογής έγινε υλοποίηση του πρωτοκόλλου BitTorrent, με βάση τις επίσημες προδιαγραφές του [26] καθώς και λαμβάνοντας υπόψη τις ιδιαίτερες συνθήκες και προϋποθέσεις που περιγράφηκαν στην προηγούμενη

ενότητα.

Το πληροφοριοκεντρικό δίκτυο που θεωρήθηκε αποτελείται από πέντε αυτόνομα συστήματα, καθένα από τα οποία έχει δεντρική μορφή με έξι επίπεδα και κάθε κόμβο να έχει 3 παιδιά, οδηγώντας σε συνολικό μέγεθος δικτύου 1820 κόμβους. Τα αυτόνομα συστήματα συνδέονται μεταξύ τους σε επίπεδο διακομιστή απομακρυσμένης ευρυζωνικής πρόσβασης (Broadband Remote Access Server - BRAS). Η επικοινωνία τους σε αυτό το επίπεδο θεωρείται ότι γίνεται πάνω από δημόσιο δίκτυο για λόγους επίδειξης.

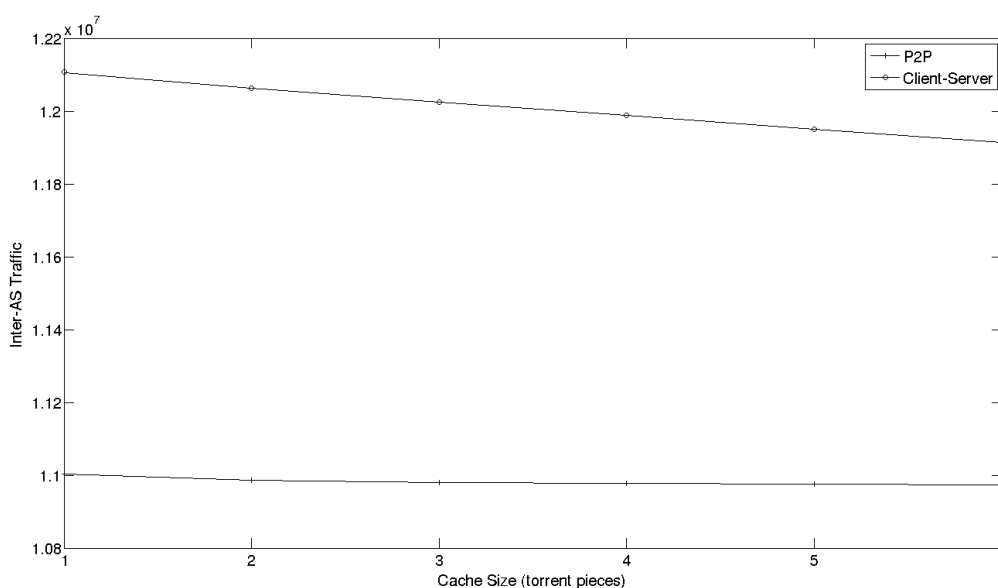
Οι παράμετροι που μετρώνται είναι η πιθανότητα εύρεσης περιεχομένου στην κρυφή μνήμη και η κίνηση που φεύγει εκτός αυτόνομου συστήματος. Ειδικά για την τελευταία, είναι προς το συμφέρον του παρόχου δικτύου να τη διατηρεί σε χαμηλά επίπεδα εφόσον παρεμβάλλεται το διαδίκτυο, καθώς η χρήση του χρεώνεται στη βάση κάποιας άδειας χρήσης. Οι δύο αυτές παράμετροι μελετώνται σε σύγκριση μεταξύ των αρχιτεκτονικών μοντέλων της ομότιμης επικοινωνίας και της επικοινωνίας πελάτη-εξυπηρετητή, η οποία υποστηρίζεται αυτή τη στιγμή από τα πληροφοριοκεντρικά δίκτυα.

Ο αριθμός των κόμβων που συμμετέχουν στο υπερκείμενο δίκτυο των χρηστών και διαμοιράζονται το περιεχόμενο ανέρχονταν στους 80.000, ομοιόμορφα κατανεμημένοι στα φύλλα των δέντρων που αντιστοιχούν στα σημεία πρόσβασης (DSLAM) των αυτόνομων συστημάτων. Οι προτιμήσεις των χρηστών σε torrent ακολουθούσαν το νόμο του Zipf με εκθετική παράμετρο 0,82 [27]. Κάθε torrent αντιστοιχεί σε ένα διαμοιραζόμενο αρχείο, τεμαχιζόμενο σε 12 κομμάτια (pieces), με κάθε κομμάτι να χωρίζεται σε 16 επιμέρους τμήματα (blocks).



Σχήμα 4: Πιθανότητα Εύρεσης Περιεχομένου στην Κρυφή Μνήμη

Η πρώτη προσέγγιση στις μετρήσεις αφορά τη συμπεριφορά των ομότιμων πρωτοκόλλων σε σύγκριση με την κλασική αρχιτεκτονική πελάτη-εξυπηρετητή όταν αλλάζει το μέγεθος της κρυφής μνήμης των κόμβων του δικτύου. Έτσι, θεωρήθηκε σταθερός αριθμός



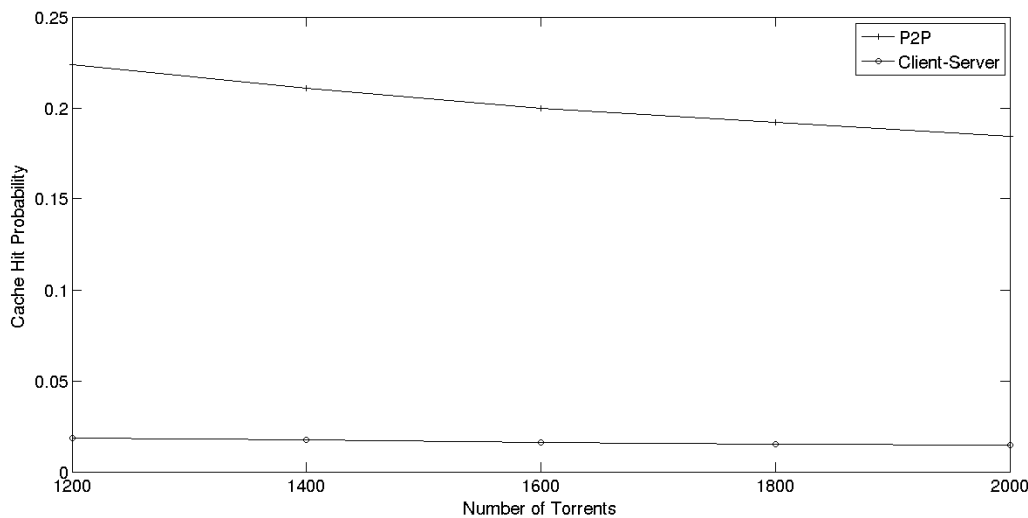
Σχήμα 5: Κίνηση εκτός Αυτόνομου Συστήματος

διαμοιραζόμενων αρχείων, ίσως με 1.000, και μέγεθος κρυφής μνήμης ανά κόμβο δικτύου από 1 μέχρι 6 κομμάτια του αρχείου. Σημειώνεται ότι θεωρήθηκαν αρχεία ίδιου μεγέθους.

Το Σχήμα 4 παρουσιάζει την εξέλιξη της πιθανότητας εύρεσης περιεχομένου στην κρυφή μνήμη για τις δύο περιπτώσεις της ομότιμης και της αρχιτεκτονικής πελάτη-εξυπηρετητή. Το πρώτο συμπέρασμα που ξεκάθαρα εξάγεται είναι ότι η πιθανότητα στην πρώτη περίπτωση είναι πολύ μεγαλύτερη από τη δεύτερη. Επιπλέον, ο ρυθμός αύξησης, αν και δείχνει γραμμικός και στις δύο περιπτώσεις, είναι επίσης μεγαλύτερος για το ομότιμο πρωτόκολλο. Συνεπώς, φαίνεται ότι τα ομότιμα πρωτόκολλα εκμεταλλεύονται με καλύτερο τρόπο την κρυφή μνήμη και, συνεπώς, τα ιδιαίτερα χαρακτηριστικά του πληροφοριοκεντρικού δικτύου.

Το όφελος αυτής της παρατήρησης αποτυπώνεται στην πράξη στο Σχήμα 5, όπου φαίνεται ο αριθμός των μηνυμάτων Interest που βγαίνουν έξω από το αυτόνομο σύστημα του αιτούμενου κόμβου. Η κίνηση αυτή έχει ως αποτέλεσμα εισερχόμενο περιεχόμενο. Όπως είναι προφανές, στην περίπτωση του ομότιμου δικτύου, η κίνηση αυτή είναι κατά πολύ μικρότερη. Φυσικά, όσο αυξάνεται το μέγεθος της κρυφής μνήμης, μειώνεται και στις δύο περιπτώσεις ο αριθμός των Interest που περνούν στο δημόσιο δίκτυο, αλλά σε κάθε περίπτωση ο αριθμός για την επικοινωνία πελάτη-εξυπηρετητή παραμένει υψηλός.

Εκτός, όμως, της επίδρασης του μεγέθους της κρυφής μνήμης, είναι ενδιαφέρον να μελετηθεί η συμπεριφορά του δικτύου με κάθε μοντέλο επικοινωνίας, όσο το περιεχόμενο γίνεται περισσότερο, πράγμα που αποτελεί άλλωστε γεγονός. Για το λόγο αυτό, πραγματοποιήθηκαν προσομοιώσεις με σταθερό μέγεθος κρυφής μνήμης και αυξανόμενο αριθμό διαμοιραζόμενων αρχείων, από 1.200 μέχρι 2.000 με βήμα 200. Και σε αυτήν την περίπτωση,



Σχήμα 6: Πιθανότητα Εύρεσης Περιεχομένου στην Κρυφή Μνήμη ως προς Αριθμό Αρχείων

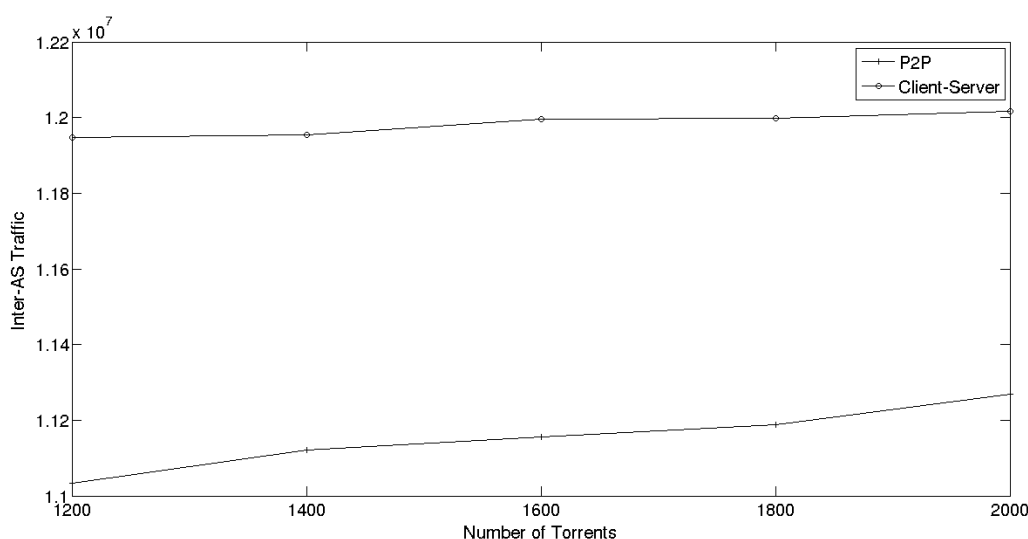
το μέγεθος των αρχείων θεωρήθηκε ίδιο για όλα.

Το Σχήμα 6 παρουσιάζει την πιθανότητα να βρεθεί το περιεχόμενο στην κρυφή μνήμη κάποιου κόμβου του δικτύου. Όπως ήταν αναμενόμενο, η πιθανότητα μειώνεται με την αύξηση του περιεχομένου. Αν και η μείωση αυτή είναι πιο αισθητή στο μοντέλο ομότιμης επικοινωνίας, η διαφορά μεταξύ των δύο παραμένει τόσο υψηλή, ώστε να μην μπορεί να αμφισβητηθεί η ανωτερότητα της απόδοσης του ομότιμου πρωτοκόλλου.

Αντίστοιχα συμπεράσματα εξάγονται και από την κίνηση έξω από το αυτόνομο σύστημα του αιτούμενου χρήστη, το οποίο παρουσιάζεται στο Σχήμα 7. Προφανώς, η κίνηση αυξάνεται με το περιεχόμενο, ωστόσο και πάλι η διαφορά ανάμεσα στην ομότιμη και την αρχιτεκτονική πελάτη-εξυπηρετητή παραμένει πολύ μεγάλη. Μάλιστα, σε αυτήν την περίπτωση, φαίνεται η αύξηση της κίνησης για τη δεύτερη περίπτωση να είναι μεγαλύτερη. Αυτό συμβαίνει διότι με τα ομότιμα πρωτόκολλα ένα κομμάτι της κίνησης μπορεί να περιορίζεται έτσι κι αλλιώς στα πλαίσια του αυτόνομου συστήματος, αφού δεν επιλέγεται μόνο ένας κόμβος εξυπηρέτησης, ο οποίος είναι κατά κανόνα σε κάποιο άλλο αυτόνομο σύστημα.

2.6 Υποστήριξη Ροής Πολυμεσικού Περιεχομένου

Το πολυμεσικό περιεχόμενο και ειδικότερα το βίντεο αναμένεται να κυριαρχήσει κατά τα επόμενα χρόνια στη χρήση του διαδικτύου. Σύμφωνα με τις προβλέψεις της CISCO [28], το έτος 2017 θα χρειάζονται πέντε εκατομμύρια χρόνια σε κάποιο χρήστη του διαδικτύου για να δει όλα τα διαθέσιμα βίντεο, με περίπου ένα εκατομμύριο λεπτά βίντεο να διατρέχουν το διαδίκτυο κάθε δευτερόλεπτο. Παγκοσμίως, το 2017 η χρήση του διαδικτύου



Σχήμα 7: Κίνηση εκτός Αυτόνομου Συστήματος ως προς Αριθμό Αρχείων

για ροή βίντεο θα αφορά το 69% της συνολικής κίνησης του διαδικτύου, αυξημένο κατά 12% από το 2012, ενώ η κίνηση για κατ'απαίτηση βίντεο (Video on Demand - VoD) αναμένεται να τριπλασιαστεί φτάνοντας τα έξι δισεκατομμύρια DVD το μήνα.

Επιπλέον αυτού, σύμφωνα με την ίδια πρόβλεψη, η χρήση των δικτύων παράδοσης περιεχομένου (Content Delivery Networks - CDNs) αναμένεται να εξυπηρετεί το 51% της συνολικής κίνησης του διαδικτύου. Ειδικά για το βίντεο, τα CDNs, αναμένεται να εξυπηρετούν το 65% της συνολικής κίνησης, αυξημένο επίσης κατά 12% σε σχέση με τις καταγραφές του 2012.

Λαμβάνοντας υπόψη τα παραπάνω, είναι προφανές ότι η επιτυχία ενός δικτύου νέας γενιάς θα δοκιμαστεί σε ένα σημαντικό βαθμό από τη δυνατότητά του να διαχειριστεί τις νέες αυτές ανάγκες που προκύπτουν. Ειδικά η αυξημένη προβλεπόμενη χρήση των CDNs, τα οποία περιορίζουν την κίνηση του διαδικτύου παρέχοντας περιεχόμενο από τις δικές τους εσωτερικές δομές δεδομένων, θυμίζοντας κατά πολύ τη λειτουργία ενός πληροφοριοκεντρικού δικτύου, αποτελεί ενθαρρυντικό παράγοντα ως προς την επιτυχία του πληροφοριοκεντρικού μοντέλου δικτυακής επικοινωνίας.

Ωστόσο, η έρευνα για την επιτυχημένη ροή βίντεο στους καταναλωτές δεν περιορίζεται μόνο στο κομμάτι του δικτύου. Τεχνολογίες πολυμέσων, οι οποίες αφορούν κωδικοποίηση (π.χ. Scalable Video Coding - SVC, Multiple Description Coding - MDC) είτε τρόπους κατεβάσματος του περιεχομένου (π.χ. Progressive Download, P2P Streaming, (Adaptive) HTTP Streaming). Ένα σύγχρονο δίκτυο θα πρέπει να λαμβάνει υπόψη αυτές τις τεχνολογίες προκειμένου να βελτιώσει την απόδοσή του.

Εκτός, όμως, της κωδικοποίησης, τα τελευταία χρόνια έχει δοθεί ιδιαίτερη έμφαση σε πρωτόκολλα προσαρμοσμένης ροής πολυμεσικού περιεχομένου με χρήση του πρωτοκόλλου HTTP (Adaptive HTTP Streaming). Διάφορες υλοποιήσεις είναι διαθέσιμες, όπως

το Apple HTTP Live Streaming [29], το Microsoft Smooth Streaming [30], το Adobe HTTP Dynamic Streaming [31] και το MPEG Dynamic Adaptive Streaming over HTTP (DASH) [32]. Αν και υπάρχουν διαφοροποιήσεις σε ό,τι αφορά την εκάστοτε υλοποίηση, η βασική ιδέα που διέπει τις προαναφερθείσες τεχνολογίες είναι ότι το βίντεο κωδικοποιείται σε διάφορες ποιότητες και διασπάται σε επιμέρους κομμάτια, καθένα από τα οποία είναι διαθέσιμο από κάποιον εξυπηρετητή πολυμέσων κάνοντας χρήση του πρωτοκόλλου HTTP. Έτσι, μια εφαρμογή αναπαραγωγής πολυμέσων, με βάση τα χαρακτηριστικά της σύνδεσης, αναλαμβάνει να ζητήσει το κατάλληλο κομμάτι του βίντεο, προκειμένου να επιτύχει τη βέλτιστη ποιότητα εμπειρίας του χρήστη.

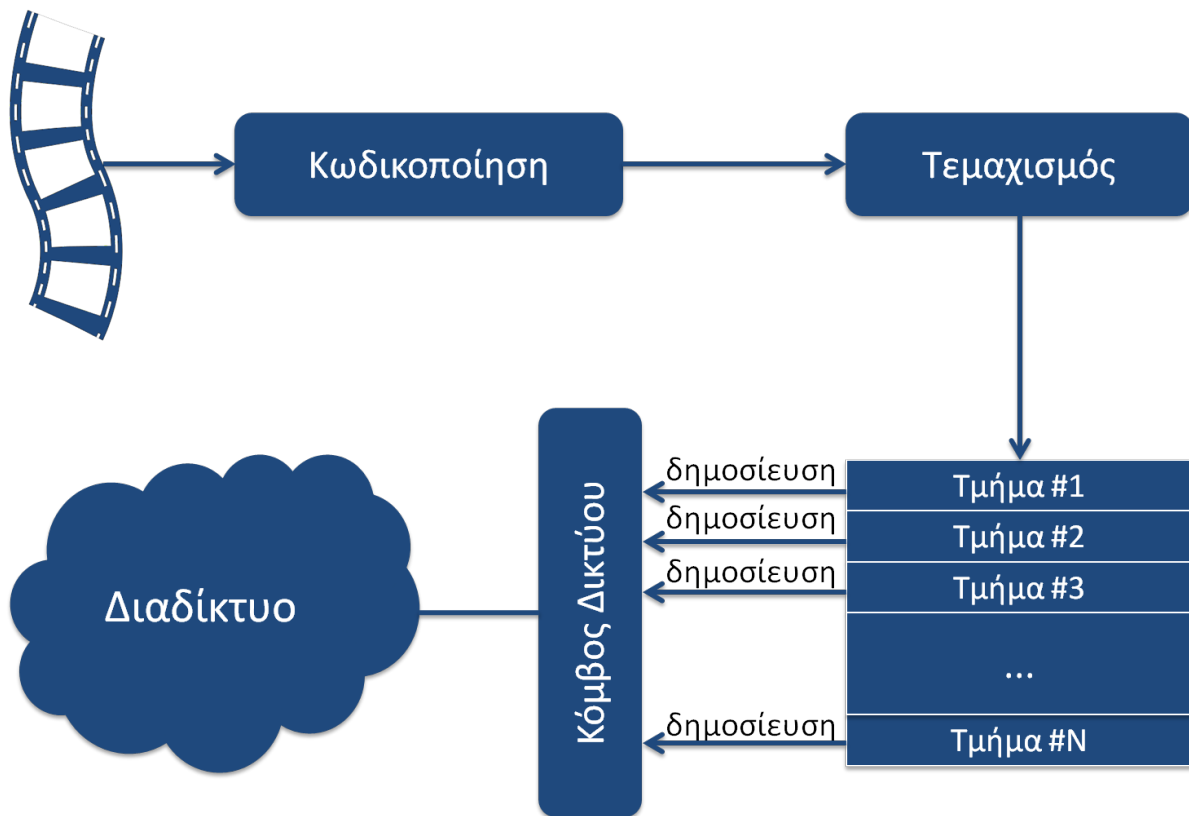
Η κινητήρια παρατήρηση για το προτεινόμενο πρωτόκολλο είναι ότι, από τη στιγμή που ένα πληροφοριοκεντρικό δίκτυο δρομολογεί τα πακέτα με βάση το όνομά τους και στην προσαρμοσμένη ροή βίντεο έχουμε τμήματα βίντεο που φέρουν διαφορετικό όνομα με βάση την ποιότητα, η απόδοση του μηχανισμού της κρυφής μνήμης θα περιορίζεται, σε σχέση με την περίπτωση της μη-προσαρμοσμένης ροής. Έτσι, με μια πρώτη προσέγγιση, ο πάροχος δικτύου θα προτιμά την τελευταία. Από την άλλη όμως, ο πελάτης θέλει όσο το δυνατόν καλύτερη ποιότητα στο βίντεο που λαμβάνει, οπότε θα προτιμούσε την πρώτη (προσαρμοσμένη) περίπτωση. Λαμβάνοντας υπόψη αυτή τη σύγκρουση συμφερόντων, το προτεινόμενο πρωτόκολλο επιχειρεί να βελτιώσει την απόδοση της προσαρμοσμένης ροής βίντεο πάνω από πληροφοριοκεντρικό δίκτυο, με οφέλη τόσο για τον πάροχο δικτύου, όσο και για τον τελικό καταναλωτή.

2.6.1 Υποστήριξη Ροής Βίντεο πάνω από Πληροφοριοκεντρικό Δίκτυο

Η αρχιτεκτονική για υποστήριξη ροής βίντεο πάνω από πληροφοριοκεντρικό δίκτυο δε διαφέρει κατά πολύ από αυτή που χρησιμοποιείται ήδη για ροή πάνω από HTTP στα υπάρχοντα δίκτυα, δεδομένου ότι και για τα τελευταία τα τμήματα του βίντεο ζητούνται με βάση το όνομά τους. Έτσι, υπάρχει ένας εξυπηρετητής ροής βίντεο, ο οποίος διαθέτει στους πελάτες τα τμήματα του περιεχομένου, με κάθε ένα από αυτά να φέρει ένα μοναδικό όνομα. Η προσέγγιση που θεωρήθηκε είναι αντίστοιχη με αυτή που έχει παρουσιαστεί από το NDN [33] καθώς και από τους συγγραφείς του [34].

Όπως περιγράφεται και στο Σχήμα 8, ο εξυπηρετητής πρώτα κωδικοποιεί το βίντεο και το τεμαχίζει σε επιμέρους τμήματα. Τα μεταδεδομένα που αφορούν αυτά τα τμήματα γράφονται σε ένα ειδικό αρχείο ευρετηρίου, το οποίο εξαρτάται από την κωδικοποίηση που χρησιμοποιείται. Στη συνέχεια, καθένα από αυτά τα τμήματα καθώς και το αντίστοιχο αρχείο ευρετηρίου δημοσιεύονται με ένα μοναδικό όνομα στο δίκτυο, μέσω του κόμβου όπου είναι συνδεδεμένος ο εξυπηρετητής.

Το όνομα που χρησιμοποιείται για κάθε τμήμα του βίντεο θα πρέπει να αντανακλά το βίντεο αυτό καθαυτό καθώς και την ποιότητα, στην οποία είναι κωδικοποιημένο. Με βάση και τις προτάσεις του NDN [33], το όνομα προτείνεται να έχει την ακόλουθη μορφή



Σχήμα 8: Δημοσίευση Βίντεο σε Πληροφοριοκεντρικό Δίκτυο

(μπορεί να είναι και πιο σύνθετη, αλλά για λόγους απλότητας, θα κρατηθεί αυτή η σημειολογία):

$L = \langle \text{VIDEO} \rangle / \langle \text{SEGMENT} \rangle / \langle \text{QUALITY} \rangle$

Το πρώτο κομμάτι του ονόματος (ετικέτα) αφορά το όνομα του βίντεο. Το δεύτερο είναι το συγκεκριμένο τμήμα του βίντεο, το οποίο δημοσιεύεται. Το τρίτο είναι η ποιότητα, στην οποία είναι κωδικοποιημένο το εν λόγω τμήμα και έχει νόημα στην περίπτωση της προσαρμοσμένης ροής. Σε περίπτωση μη-προσαρμοσμένης ροής, το τελευταίο πεδίο μπορεί να παραληφθεί.

Στην πλευρά του πελάτη, υποθέτουμε ότι η εφαρμογή αναπαραγωγής του βίντεο πραγματοποιεί αιτήσεις, όπως ακριβώς θα έκανε και στην περίπτωση της ροής υποστηριζόμενης από HTTP. Σημειώνεται ότι δεν είναι απαραίτητο να δημιουργηθεί μια νέα εφαρμογή αναπαραγωγής που να χρησιμοποιεί το πληροφοριοκεντρικό δίκτυο. Αντιθέτως, αρκεί να χρησιμοποιηθεί κάποιος πληρεξούσιος (proxy) HTTP, ο οποίος αναλαμβάνει να μετατρέψει τις αιτήσεις και τις απαντήσεις στη μορφή που ήδη καταλαβαίνει η εφαρμογή αναπαραγωγής. Αυτή η προσέγγιση ακολουθήθηκε άλλωστε και στο [34].

Κάθε αίτηση του πελάτη μεταφράζεται σε ένα μήνυμα Interest του δικτύου, το οποίο δρομολογείται προς τον εξυπηρετητή βίντεο. Κάθε ενδιαμέσος δρομολογητής που λαμβά-

νει αυτό το μήνυμα ελέγχει την κρυφή του μνήμη, προκειμένου να εξυπηρετήσει την αίτηση τοπικά. Γενικά, είναι προς το συμφέρον του εκάστοτε παρόχου δικτύου να εξυπηρετεί τις αιτήσεις τοπικά, ώστε να μειώσει τη χρήση του διαδικτύου. Συνεπώς, η πιθανότητα να βρει το περιεχόμενο στην κρυφή μνήμη έχει ιδιαίτερη σημασία για αυτόν. Άλλωστε, αυτός είναι και ο βασικός λόγος της επιτυχίας των δικτύων παράδοσης περιεχομένου, παρά το υψηλό τους κόστος για τους παρόχους.

2.6.2 Εγγενής Υποστήριξη Προσαρμοσμένης Ροής Βίντεο στο Δίκτυο

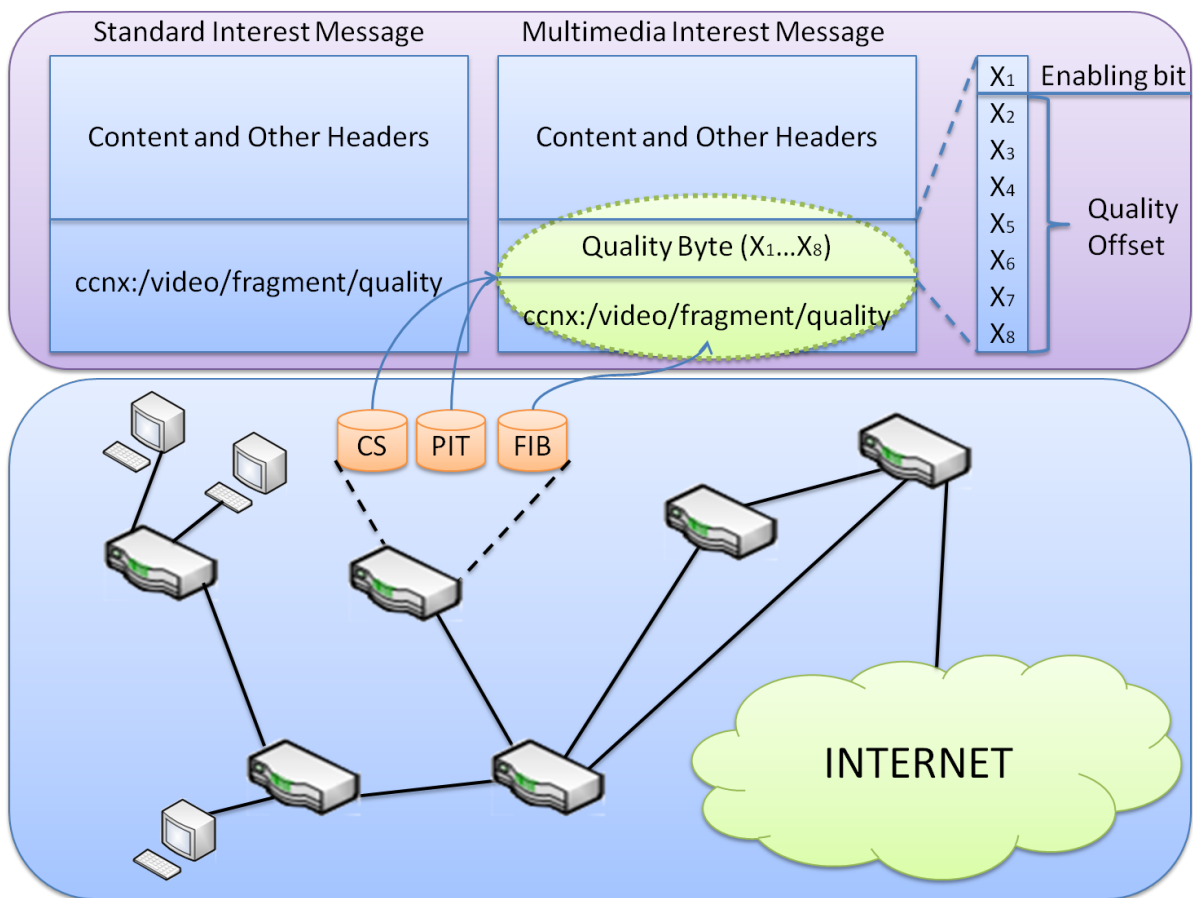
Οι προτεινόμενες επεκτάσεις του πληροφοριοκεντρικού δικτύου και πιο συγκεκριμένα του μοντέλου του CCNx που χρησιμοποιήθηκε για την εν λόγω προσέγγιση, στοχεύει στη βελτίωση τόσο της πλευράς του χρήστη με καλύτερη ποιότητα λαμβανόμενου βίντεο, όσο και του παρόχου υπηρεσιών διαδικτύου με υψηλότερη πιθανότητα οι αιτήσεις των χρηστών να εξυπηρετούνται από τους τοπικούς δρομολογητές. Η βασική ιδέα είναι ότι το δίκτυο χρησιμοποιεί τα ονόματα των τμημάτων του εκάστοτε βίντεο όχι σαν απλές ετικέτες, αλλά πηγαίνει ένα βήμα παραπέρα λαμβάνοντας υπόψη και τη σημασιολογία του ονόματος.

Στο CCNx, όταν ένας κόμβος του δικτύου παραλάβει κάποιο μήνυμα Interest, τότε θα ελέγξει πρώτα την προσωρινή μνήμη του για περιεχόμενο με το όνομα της αίτησης, ώστε να την εξυπηρετήσει τοπικά. Στην περίπτωση της προσαρμοσμένης ροής βίντεο, το όνομα αυτό θεωρούμε ότι έχει τη μορφή που παρουσιάστηκε στην προηγούμενη ενότητα. Για παράδειγμα, αν θεωρήσουμε ότι έγινε μια αίτηση για το δεύτερο τμήμα της ταινίας casablanca σε ποιότητα υψηλής ευκρίνειας, τότε το όνομα θα έχει την ακόλουθη μορφή:

```
ccnx:/casablanca/2/high-definition
```

Ένα ζήτημα που προκύπτει εδώ είναι ότι ο κόμβος του δικτύου μπορεί να διατηρεί στη μνήμη του το τμήμα του βίντεο που αιτείται ο πελάτης, αλλά σε διαφορετική ποιότητα. Σε μια τέτοια περίπτωση, στην τυπική υλοποίηση του CCNx, δε θα μπορέσει να εξυπηρετηθεί η αίτηση τοπικά. Στην προτεινόμενη αρχιτεκτονική, ο κόμβος θα απαντήσει με την ποιότητα που ήδη διαθέτει, ακολουθώντας λογική βέλτιστης προσπάθειας (best effort). Με αυτόν τον τρόπο, όπως φαίνεται και στην Ενότητα 2.6.3, επωφελούνται τόσο ο πάροχος υπηρεσιών διαδικτύου όσο και ο καταναλωτής του βίντεο.

Για να μπορέσει ο δρομολογητής να ακολουθήσει αυτήν την προσέγγιση, χρειάστηκε να επεκταθεί το τυπικό μήνυμα Interest του CCNx, στο *Multimedia Interest*, το οποίο, όπως φαίνεται και στο Σχήμα 9, περιλαμβάνει ένα επιπλέον byte πληροφορίας, το *byte ποιότητας*. Το πρώτο (πιο σημαντικό) bit του byte ποιότητας ονομάζεται *bit ενεργοποίησης* και υποδεικνύει εάν ο κόμβος δικτύου που λαμβάνει το Interest επιτρέπεται να κάνει ταίριασμα βέλτιστης προσπάθειας. Τα τελευταία 7 bits του byte ποιότητας υποδεικνύουν το πλήθος



Σχήμα 9: Υποστήριξη Ροής Βίντεο στο Δίκτυο

των χαρακτήρων του ονόματος που αφορούν την ποιότητα, ώστε να μπορέσει ο δρομολογητής να ξεχωρίσει το κομμάτι που αφορά το βίντεο και το τμήμα, από το κομμάτι που αφορά την ποιότητα. Τα 7 bits επιτρέπουν μήκος 127 χαρακτήρων για την ποιότητα στο όνομα, το οποίο θεωρείται επαρκές.

Ακολουθώντας αυτήν την προσέγγιση, όταν ο δρομολογητής λάβει ένα μήνυμα Multimedia Interest, εξάγει πρώτα το όνομα και το byte ποιότητας. Αν το bit ενεργοποίησης έχει τιμή 1, τότε ο δρομολογητής θα ψάξει στην τοπική του μνήμη για κάποιο τμήμα του βίντεο ακόμα και σε διαφορετική ποιότητα και, αν το βρει, επιστρέφει εκείνο. Η αναζήτηση γίνεται με ταίριασμα προθέματος του ονόματος του βίντεο, έχοντας αφαιρέσει τους χαρακτήρες που υποδεικνύουν την ποιότητα, όπως καθορίζεται από το byte ποιότητας.

Στην περίπτωση που κάποιο μήνυμα Multimedia Interest εξυπηρετηθεί από διαφορετική ποιότητα από τη ρητώς αιτούμενη, ανακύπτει ζήτημα για τον τρόπο που θα γίνει η δρομολόγηση πίσω στον πελάτη. Ο τρόπος λειτουργίας του CCNx υπαγορεύει ότι το περιεχόμενο ακολουθεί το ίδιο μονοπάτι με την αντίστοιχη αίτηση. Για να επιτευχθεί αυτό, κάθε ενδιαμέσος δρομολογητής αντιστοιχίζει στον πίνακα PIT το μήνυμα Interest με τη διεπαφή δικτύου, από την οποία το λαμβάνει. Ακολουθώντας την προτεινόμενη προσέγγιση

όμως, το όνομα που ζήτησε ο πελάτης διαφέρει από αυτό της απάντησης στο πεδίο της ποιότητας. Για το λόγο αυτό, και μόνο στην περίπτωση που δίνεται ως απάντηση διαφορετική ποιότητα, το μήνυμα Data φέρει επιπλέον πεδίο με το όνομα που αρχικά αιτήθηκε ο πελάτης, προκειμένου να χρησιμοποιηθεί εκείνο για την προώθηση του τμήματος βίντεο πίσω στον αιτούμενο χρήστη.

Η προτεινόμενη αρχιτεκτονική προϋποθέτει ότι ο κάθε κόμβος του δικτύου είναι σε θέση να πραγματοποιεί ταίριασμα προθέματος στο επίπεδο του PIT και του Content Store (τοπική κρυφή μνήμη). Για το σκοπό αυτό μπορούν να χρησιμοποιηθούν υπάρχοντες λύσεις που ήδη χρησιμοποιούνται στο NDN [35], αλλά και εναλλακτικές λύσεις όπως τα ιεραρχικά φίλτρα Bloom [36], τα οποία είναι ιδιαίτερα αποδοτικά στις λειτουργίες ταϊριάσματος προθέματος.

Ο Αλγόριθμος 1 περιγράφει τη διαδικασία που ακολουθείται σε ένα κόμβο του δικτύου όταν φτάνει ένα μήνυμα Multimedia Interest. Αρχικά το byte ποιότητας εξάγεται από το μήνυμα και ελέγχεται το πιο σημαντικό bit (MSB). Σε περίπτωση που είναι ενεργοποιημένο, τότε τα τελευταία 7 bits χρησιμοποιούνται για να εξαχθεί από το όνομα το πρόθεμα, το οποίο θα εφαρμοστεί στη διαδικασία ταϊριάσματος προθέματος στο Content Store και στο PIT, ανάλογα με την περίπτωση. Αν το MSB δεν είναι ενεργοποιημένο, τότε ακολουθείται η τυπική διαδικασία. Στο τέλος της διαδικασίας, επιστρέφεται το περιεχόμενο της δομής δεδομένων όπου έγινε η αναζήτηση.

Αλγόριθμος 1: Εξυπηρέτηση Μηνύματος Multimedia Interest σε PIT και Content Store

Data: x είναι ένα μήνυμα Multimedia Interest που καταφθάνει σε ένα κόμβο του δικτύου

Result: το περιεχόμενο (από Content Store ή PIT) που έχει αντιστοιχηθεί το μήνυμα Multimedia Interest

```

1  $q \leftarrow x.qualityByte;$ 
2  $Content \leftarrow \emptyset;$ 
3 if  $q \& (1 \ll 7)$  then // MSB == 1
4    $offset \leftarrow q \& 127;$ 
5    $namePrefix \leftarrow x.name \gg offset;$ 
6    $Content \leftarrow hash.get(namePrefix*);$ 
7 else // MSB == 0
8    $Content \leftarrow hash.get(x.name);$ 
9 end
10 return  $Content;$ 

```

2.6.3 Πειραματική Εκτίμηση Επιδόσεων Πρωτοκόλλων

2.6.3.1 Στρατηγική Πειραμάτων

Η εκτίμηση της απόδοσης των πρωτοκόλλων ροής βίντεο έγινε με βάση τον προσομοιωτή JCCNxSim, όπως παρουσιάστηκε νωρίτερα στην Ενότητα 2.4. Η τοπολογία που

χρησιμοποιήθηκε για την αναπαράσταση του δικτύου είναι δενδρική, ακολουθώντας ανάλογες προσεγγίσεις της βιβλιογραφίας, όπως στο [37]. Εντούτοις, αντί για δυαδικό, θεωρήθηκε δένδρο με 6 επίπεδα όπου κάθε κόμβος έχει 3 παιδιά να αντιπροσωπεύει ένα αυτόνομο σύστημα κάποιου παρόχου υπηρεσιών διαδικτύου. Η ρίζα του δένδρου αναπαριστά το διακομιστή απομακρυσμένης ευρυζωνικής πρόσβασης (Broadband Remote Access Server - BRAS), τα φύλλα τους πολυπλέκτες πρόσβασης των συνδρομητικών ψηφιακών γραμμών (Digital Subscriber Line Access Multiplexer - DSLAM), ενώ οι ενδιάμεσοι κόμβοι αναπαριστούν τα δίκτυο μετάδοσης (Transport Network - TN), το οποίο συνήθως υλοποιείται με τεχνολογίες ATM ή Ethernet.

Για τη διεξαγωγή των προσομοιώσεων χρησιμοποιήθηκαν διάφορες παράμετροι για τα χαρακτηριστικά του δικτύου. Έτσι, το δίκτυο μετάδοσης θεωρήθηκε ότι υποστηρίζει ρυθμούς μετάδοσης της τάξης των 155Mbps, που αποτελεί τυπική ταχύτητα για δίκτυα ATM. Οι συνδρομητές συνδέονται στο δίκτυο με ταχύτητες 2, 8 και 16Mbps, ανάλογα με την περίπτωση που μελετάται κάθε φορά. Επιπλέον, θεωρήθηκε κυμαινόμενη κίνηση της τάξης του 30% του συνολικού εύρους ζώνης και μέχρι το 1Mbps για άλλες εφαρμογές που απαιτούν χρήση του διαδικτύου, όπως για παράδειγμα εφαρμογές ομοτίμων δικτύων ή ακόμα και απλή χρήση του παγκόσμιου ιστού.

Στην πλευρά του πελάτη υλοποιήθηκε ένας προσομοιωτής εφαρμογής αναπαραγωγής βίντεο, ο οποίος κάθε φορά που λαμβάνει ένα τμήμα του βίντεο, πραγματοποιεί μια αίτηση για το επόμενο. Ειδικά για την περίπτωση της προσαρμοσμένης ροής βίντεο, η εφαρμογή αναπαραγωγής πρέπει να υλοποιεί και μια πολιτική που υπαγορεύει πότε θα πρέπει να γίνει αλλαγή ποιότητας σε μεγαλύτερη ή μικρότερη. Μετά από σειρά πειραμάτων που αφορούσαν τα σενάρια που παρουσιάζονται στην παρούσα διατριβή και με γνώμονα ότι στην προσαρμοσμένη ροή είναι ανεπιθύμητοι οι ενδιάμεσοι ανενεργοί χρόνοι κατά την αναπαραγωγή (πάγωμα) του βίντεο, η πολιτική που ακολουθήθηκε είναι η εξής: όταν ο ενταμιευτής της εφαρμογής έχει μόνο ένα τμήμα του βίντεο, τότε η ποιότητα πέφτει κατά μία μονάδα στην κλίμακα, ενώ όταν έχει δώδεκα τμήματα, τότε ανεβαίνει κατά μία μονάδα. Επιπλέον, εισήχθη παράμετρος ευαισθησίας της τάξης του 10%, η οποία προστίθεται στα παραπάνω όρια για κάθε επίπεδο ποιότητας, καθώς όσο αυξάνεται η τελευταία, τόσο επιβαρύνεται το δίκτυο και η εφαρμογή αναπαραγωγής πρέπει να γίνεται πιο προσεκτική στις μεταβάσεις, ώστε να μην πέσει σε ατέρμονες βρόχους σκαμπαναβεσμάτων ανάμεσα σε δύο ποιότητες.

Δεδομένου ότι ο ισχυρισμός πίσω από την προτεινόμενη αρχιτεκτονική είναι ότι παρουσιάζει πλεονεκτήματα τόσο για τον πάροχο υπηρεσιών διαδικτύου όσο και για τον πελάτη, μελετώνται η πιθανότητα εύρεσης περιεχομένου στην κρυφή μνήμη των κόμβων του αυτόνομου συστήματος όπου συνδέεται ο πελάτης καθώς και η ποιότητα βίντεο που φτάνει στον τελευταίο. Ειδικά για την ποιότητα του βίντεο είναι γενικά επιθυμητό να είναι, αφενός, υψηλή και, αφετέρου, σταθερή, δεδομένου ότι τα σκαμπανεβάσματα δημιουργούν κακή εμπειρία χρήστη. Στις προσομοιώσεις θεωρήθηκε ότι ο πελάτης επιθυμεί να δει ολό-

κληρο το βίντεο χωρίς διακοπές και συνάμα με την ποιότητα μετρήθηκαν επίσης και οι ανενεργοί χρόνοι τόσο μέχρι να ξεκινήσει να παίζει το βίντεο όσο και κατά τη διάρκεια της αναπαραγωγής, συνολικά ακολουθώντας μια προσέγγιση ανάλογη με αντίστοιχες εργασίες στη βιβλιογραφία ([38][39]).

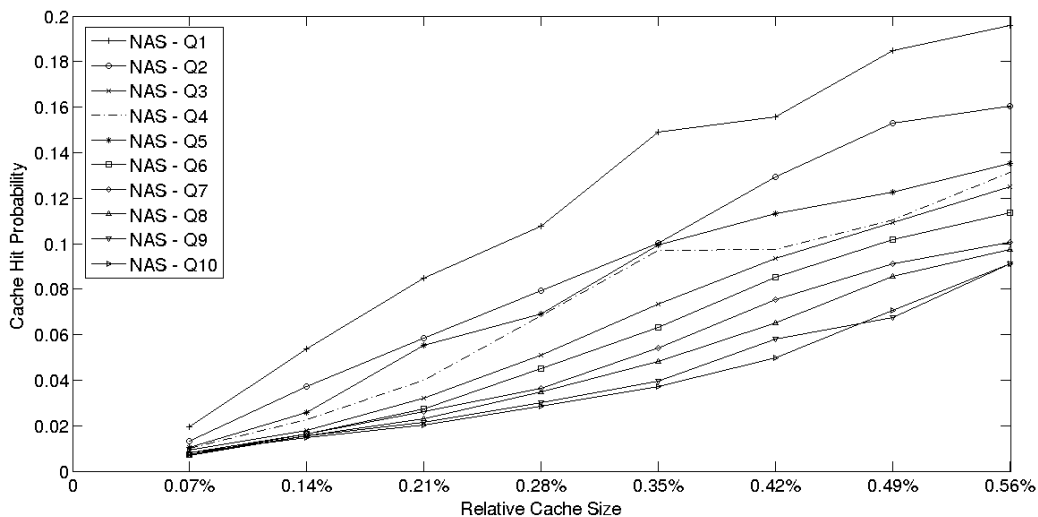
Το σύστημα που προσομοιώθηκε απαρτίζεται από 5000 κόμβους πελάτες, καθένας από τους οποίους επιλέγει 1 από 100 διαθέσιμα βίντεο διάρκειας 30 λεπτών έκαστο. Η διάρκεια αναπαραγωγής του κάθε τμήματος είναι 10 δευτερόλεπτα, ακολουθώντας την πρόταση του HTTP Live Streaming [29]. Οι προτιμήσεις των χρηστών σε βίντεο ακολουθούν το νόμο του Zipf, όπως προτείνεται στη βιβλιογραφία [27]. Το μέγεθος της κρυφής μνήμης του κάθε κόμβου του δικτύου θεωρήθηκε ως ποσοστό του συνολικού περιεχομένου που πρόκειται να διακινηθεί για βίντεο. Για τη γραμμή των 2Mbps θεωρήθηκε ελάχιστη ποιότητα βίντεο στα 400kbps, ενώ διατίθενται άλλες 10 ανώτερες ποιότητες, με καθεμιά από αυτές να είναι κατά 200kbps καλύτερη από την προηγούμενη. Η κρυφή μνήμη για το βίντεο κυμαίνεται από 0.7% ως 0.56% με βήμα 0.7% του συνολικού περιεχομένου στην ελάχιστη ποιότητα. Για να γίνει καλύτερη σύγκριση μεταξύ της τυπικής προσαρμοσμένης ροής με τις επεκτάσεις που παρουσιάστηκαν νωρίτερα στην Ενότητα 2.6.2, θεωρήθηκαν μεγαλύτερα μεγέθη γραμμής, κρυφής μνήμης καθώς και ανώτερες ποιότητες βίντεο. Έτσι, σε αυτές τις περιπτώσεις θεωρήθηκε βασική ποιότητα 1000kbps με κάθε ανώτερη ποιότητα να είναι 400kbps καλύτερη από την προηγούμενη και μέγεθος κρυφής μνήμης ανά κόμβο 0.25% ως 2% με βήμα 0.25% του συνολικού περιεχομένου στην ελάχιστη ποιότητα.

2.6.3.2 Ανάλυση Αποτελεσμάτων - Αποψη Πλευράς Παρόχου Δικτύου

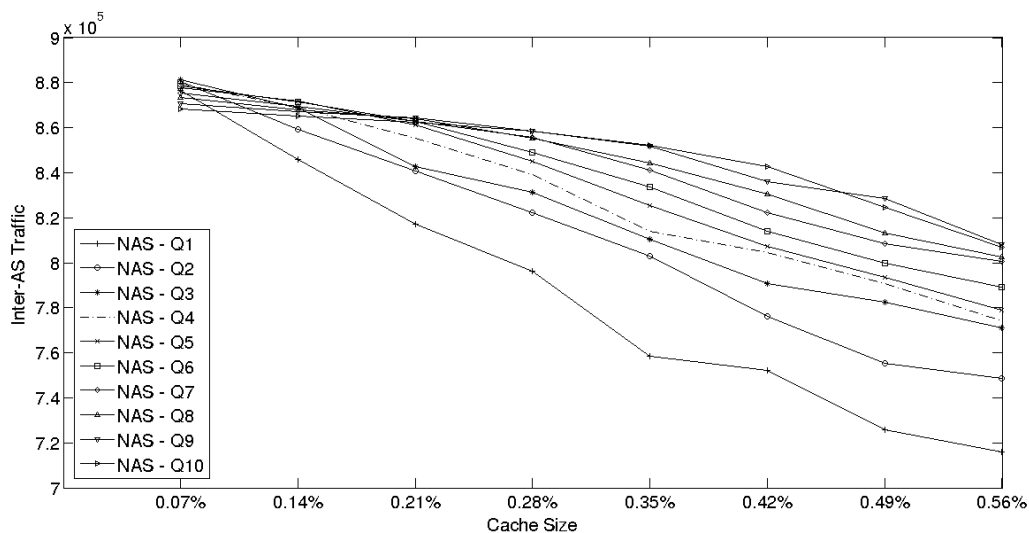
Στόχος του παρόχου δικτύου είναι να ελαχιστοποιήσει την κίνηση που φεύγει από το δίκτυό του προς το διαδίκτυο. Ένα από τα μεγαλύτερα πλεονεκτήματα των πληροφοριοκεντρικών δικτύων, και ιδιαίτερα της δρομολόγησης με βάση το όνομα, είναι ότι επιτρέπει τη χρήση της κρυφής μνήμης σε κάθε δρομολογητή, προκειμένου ο τελευταίος να είναι σε θέση να εξυπηρετήσει κάποια αίτηση τοπικά. Αυτό έχει ως αποτέλεσμα τη μείωση της κίνησης που φεύγει προς το δημόσιο διαδίκτυο, η χρήση του οποίου για κάποιον πάροχο γίνεται στη βάση κάποιας σύμβασης (Service License Agreement - SLA) που συνοδεύεται και από αντίστοιχες χρεώσεις για τον πάροχο.

Το Σχήμα 10 παρουσιάζει την πιθανότητα εύρεσης περιεχομένου στην κρυφή μνήμη κάποιου δρομολογητή του αυτόνομου συστήματος όπου είναι συνδεδεμένος ο πελάτης, σε σχέση με το μέγεθος της μνήμης για την περίπτωση της μη-προσαρμοσμένης ροής βίντεο. Όπως φαίνεται και στο σχήμα, η πιθανότητα αυτή αυξάνεται με το μέγεθος της μνήμης, σχεδόν γραμμικά, ενώ, όπως είναι φυσικό, μειώνεται όσο ανεβαίνει η ποιότητα, καθώς το μέγεθος του βίντεο είναι μεγαλύτερο, ενώ το μέγεθος της μνήμης παραμένει σταθερό και σχετικό με την ελάχιστη ποιότητα.

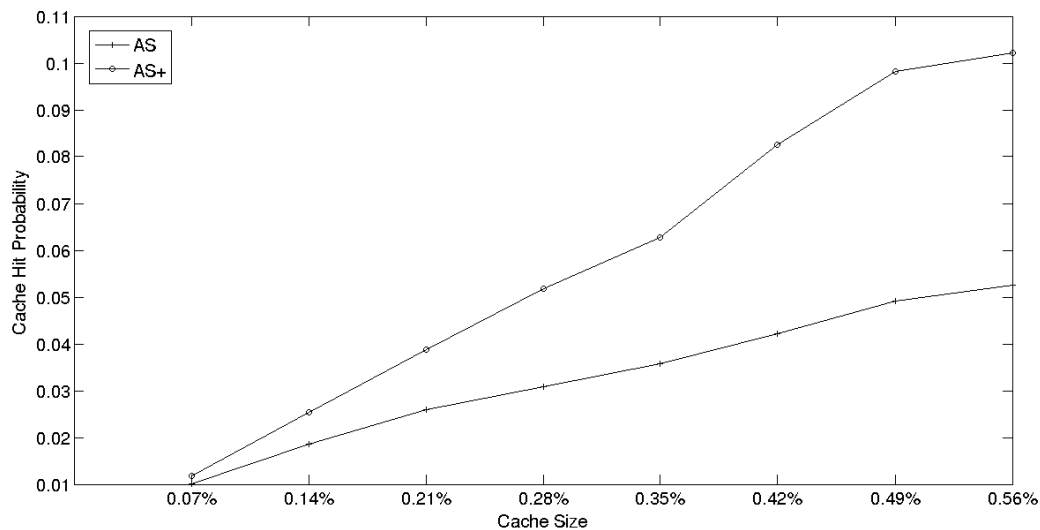
Αντίστοιχα, το Σχήμα 11 παρουσιάζει την κίνηση (σε μηνύματα Interest) που βγαί-



Σχήμα 10: Πιθανότητα Εύρεσης Περιεχομένου στην Κρυφή Μνήμη για Μη-Προσαρμοσμένη Ροή Βίντεο



Σχήμα 11: Κίνηση Έξω από το Αυτόνομο Σύστημα για Μη-Προσαρμοσμένη Ροή Βίντεο



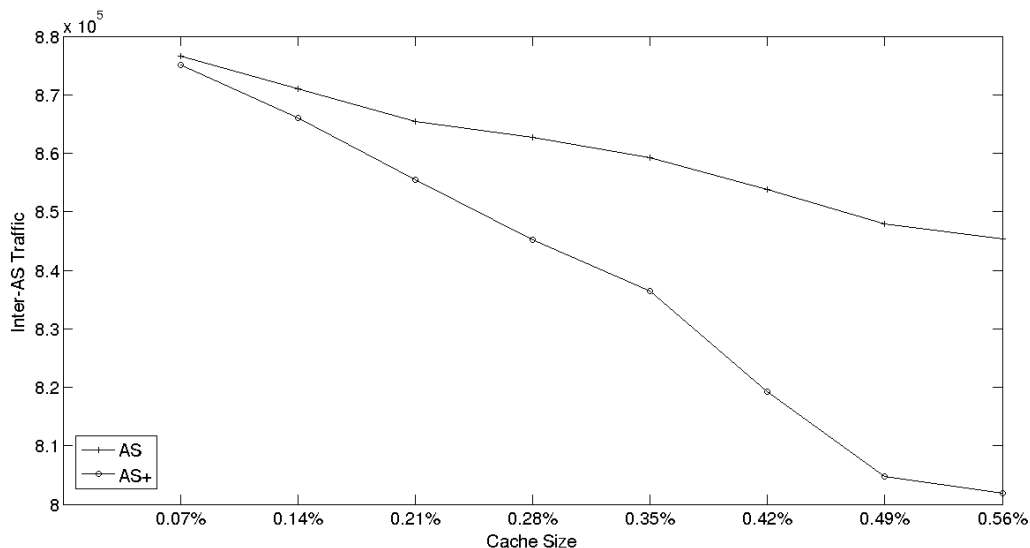
Σχήμα 12: Πιθανότητα Εύρεσης Περιεχομένου στην Κρυφή Μνήμη για Προσαρμοσμένη Ροή Βίντεο

νουν έξω από το αυτόνομο σύστημα και, στη θεώρηση που έχει γίνει, έξω από το δίκτυο του παρόχου. Όπως ήταν αναμενόμενο, λαμβάνοντας υπόψη και το προηγούμενο σχήμα με την πιθανότητα εύρεσης περιεχομένου στην κρυφή μνήμη, η κίνηση μειώνεται με την αύξηση του μεγέθους της μνήμης και αυξάνεται με την αύξηση της ποιότητας.

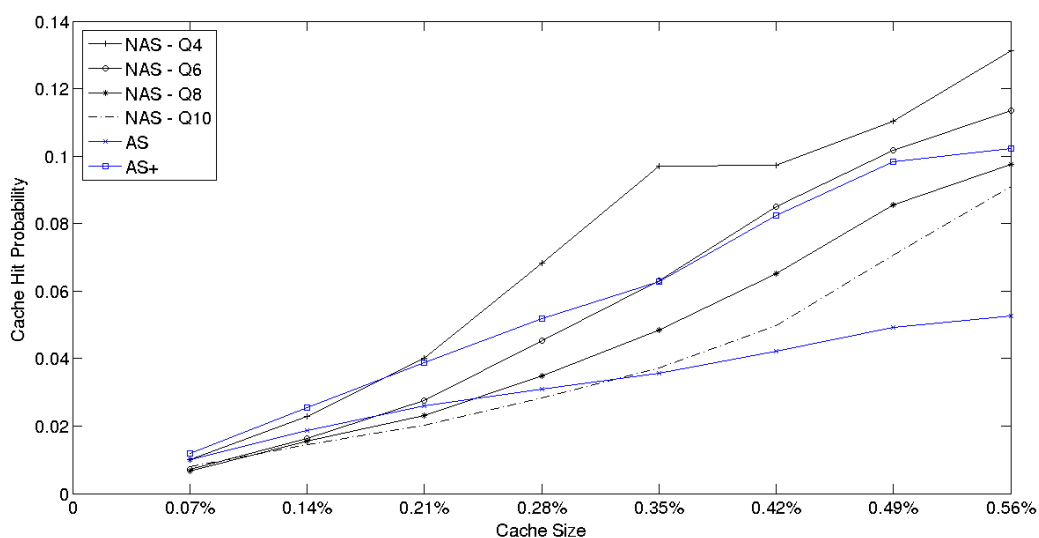
Προχωρώντας στην περίπτωση της προσαρμοσμένης ροής βίντεο, το Σχήμα 12 παρουσιάζει την πιθανότητα να βρεθεί το αιτούμενο περιεχόμενο στην κρυφή μνήμη κάποιου δρομολογητή του δικτύου στις περιπτώσεις της τυπικής αρχιτεκτονικής ροής βίντεο (AS) και της αρχιτεκτονικής που περιγράφηκε νωρίτερα με την επέκταση των μηνυμάτων Interest σε Multimedia Interest (AS+). Όπως είναι προφανές από το σχήμα, η πιθανότητα όχι μόνο είναι μεγαλύτερη στην περίπτωση της επέκτασης, αλλά η διαφορά μεταξύ των δύο μεγαλώνει με την αύξηση του μεγέθους της κρυφής μνήμης. Αυτό οφείλεται στο ότι με την επέκταση γίνεται εξυπνότερη χρήση της κρυφής μνήμης, υπό την έννοια ότι δεν υπάρχουν επικαλύψεις ίδιων τμημάτων βίντεο σε διαφορετική ποιότητα.

Το αποτέλεσμα αυτής της συμπεριφοράς αντανάκλαται στο Σχήμα 13 όπου αναπαρίσταται η κίνηση έξω από το αυτόνομο σύστημα του παρόχου δικτύου προς το διαδίκτυο. Όμοια και με την περίπτωση της μη-προσαρμοσμένης ροής βίντεο, η κίνηση αυτή ακολουθεί την αντίστροφη πορεία της πιθανότητα εύρεσης περιεχομένου στην κρυφή μνήμη. Έτσι, ο πάροχος δικτύου φαίνεται ότι αποκομίζει σημαντικά οφέλη από την εφαρμογή της επέκτασης στην προσαρμοσμένη ροή βίντεο, σε σχέση με τη μη-προσαρμοσμένη.

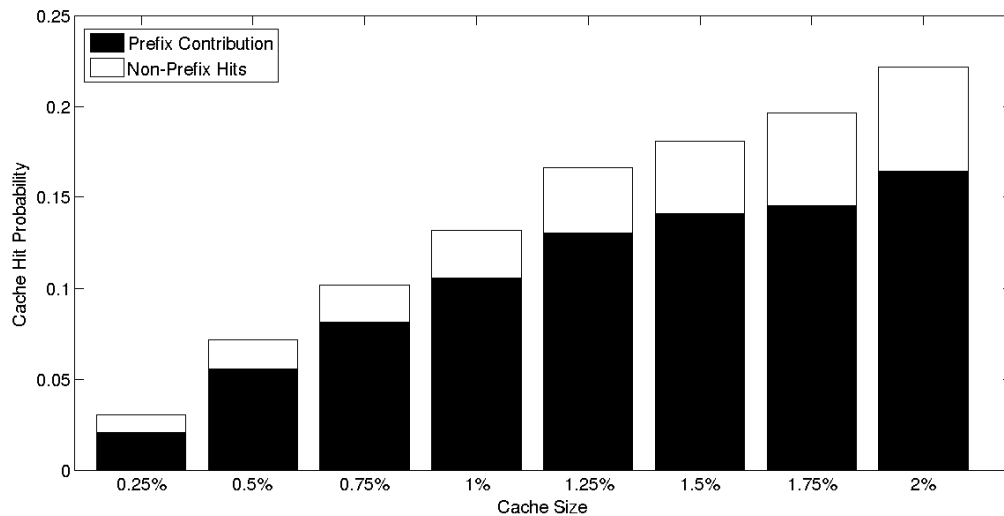
Συγκρίνοντας και τις τρεις προσεγγίσεις, μη-προσαρμοσμένη, προσαρμοσμένη και προσαρμοσμένη με την επέκταση ροή βίντεο, όπως φαίνεται στο Σχήμα 14, είναι γενικά προς το συμφέρον του παρόχου δικτύου να γίνονται αιτήσεις για βίντεο ελάχιστης ποιότητας. Παρόλα αυτά, όσο αυξάνεται η ποιότητα και συγκεκριμένα από τα 1200kbps μέχρι



Σχήμα 13: Κίνηση Έξω από το Αυτόνομο Σύστημα για Προσαρμοσμένη Ροή Βίντεο



Σχήμα 14: Σύγκριση Προσεγγίσεων ως προς την Πιθανότητα Εύρεσης Περιεχομένου στην Κρυφή Μνήμη



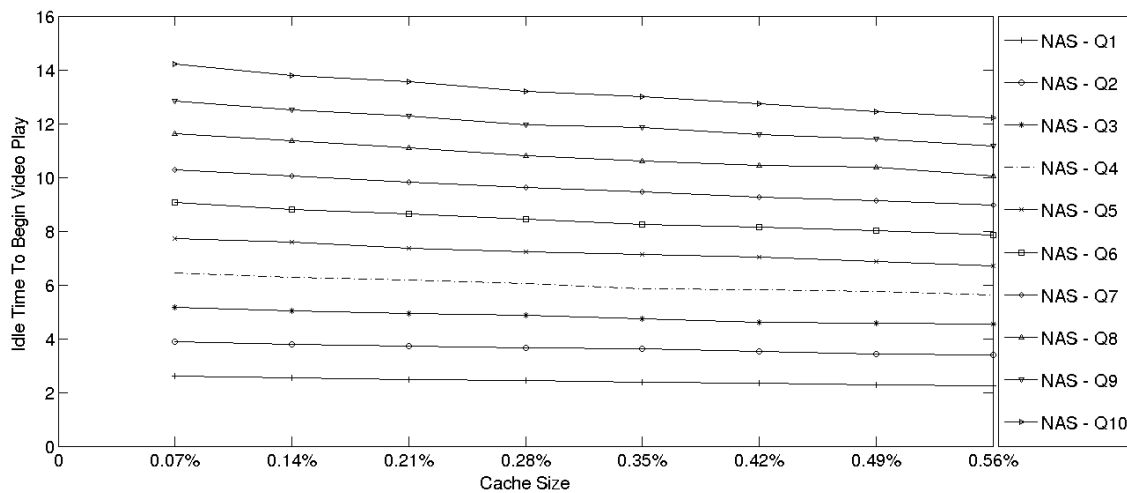
Σχήμα 15: Επιρροή Επέκτασης στην Πιθανότητα Εύρεσης Περιεχομένου στην Κρυφή Μνήμη

τα 1600kbps, φαίνεται ότι η συμπεριφορά είναι αντίστοιχη με την προσαρμοσμένη ροή βίντεο με την προτεινόμενη επέκταση. Συνεπώς, αν είναι και προς το συμφέρον του τελικού χρήστη, ο πάροχος δεν επηρεάζεται αρνητικά εφόσον ο πρώτος αιτείται προσαρμοσμένης ροής σε σχέση με σταθερή ποιότητα μεγαλύτερη των 1600kbps. Η κατάσταση για την τυπική περίπτωση προσαρμοσμένης ροής είναι χειρότερη για τον πάροχο, καθώς για μεγαλύτερα μεγέθη μνήμης η απόδοση της κρυφής μνήμης είναι χειρότερη και από τη βέλτιστη ποιότητα σταθερής ροής, δηλαδή τα 2400kbps.

Η απόδοση της προτεινόμενης επέκτασης για τον πάροχο δικτύου στην προσαρμοσμένη ροή βίντεο αντικατοπτρίζεται στο Σχήμα 15 όπου φαίνεται το ποσοστό εύρεσης περιεχομένου στις κρυφές μνήμες των δρομολογητών με χρήση του προθέματος. Το συγκεκριμένο γράφημα αφορά μεγαλύτερη γραμμή της τάξης των 16Mbps, υψηλότερες ποιότητες βίντεο, ξεκινώντας από τα 1000kbps και μεγαλύτερα μεγέθη κρυφής μνήμης, προκειμένου να είναι πιο εμφανής η επίδραση της επέκτασης. Έτσι, όπως είναι φανερό και από το σχήμα, το μεγαλύτερο ποσοστό των επιτυχιών στην κρυφή μνήμη γίνεται με τη χρήση προθέματος, πράγμα που αποδεικνύει το όφελος που υπάρχει από τη συγκεκριμένη προσέγγιση για τον πάροχο υπηρεσιών διαδικτύου.

2.6.3.3 Ανάλυση Αποτελεσμάτων - Αποψη Πλευράς Τελικού Χρήστη

Εξετάζοντας το πρωτόκολλο από τη σκοπιά του τελικού χρήστη, το βάρος δόθηκε σε τρεις βασικές παραμέτρους: τον ανενεργό χρόνο από τη στιγμή που ο χρήστης αιτείται να ξεκινήσει η αναπαραγωγή του βίντεο μέχρι αυτή να ξεκινήσει, ο ανενεργός χρόνος κατά τη διάρκεια της αναπαραγωγής (πάγωμα) του βίντεο και η ποιότητα που φτάνει τελικά στην εφαρμογή αναπαραγωγής. Ειδικά για την περίπτωση της προσαρμοσμένης



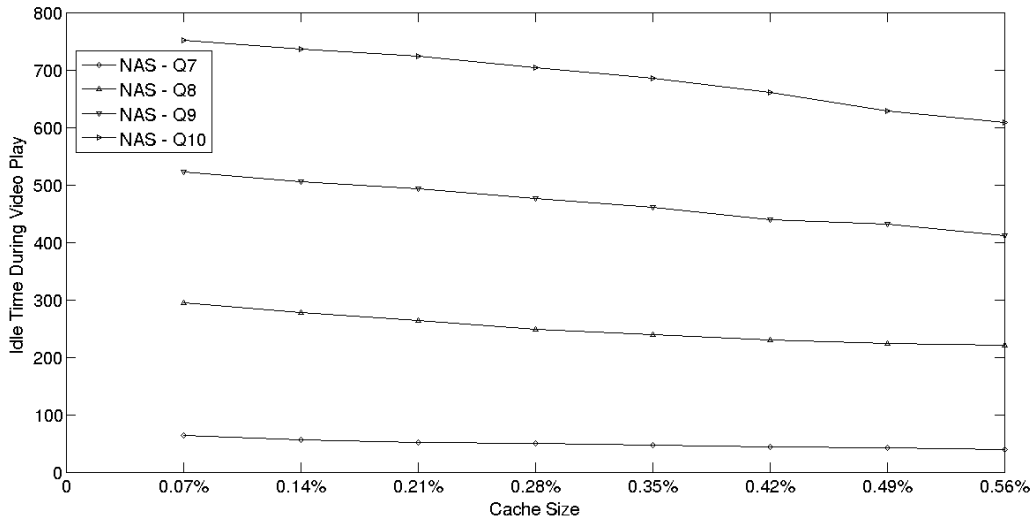
Σχήμα 16: Χρόνος Αναμονής μέχρι την Εκκίνηση Αναπαραγωγής στη Μη-Προσαρμοσμένη Ροή

ροής, κατόπιν αριθμού πειραμάτων, καθορίστηκαν οι κατάλληλες παράμετροι έτσι ώστε να εκμηδενίζονται οι ανενεργοί χρόνοι κατά τη διάρκεια αναπαραγωγής του βίντεο, καθώς θεωρήθηκε ότι η επιλογή ενός προσαρμόσιμου πρωτοκόλλου γίνεται κυρίως για να αποφευχθούν τέτοιου είδους φαινόμενα.

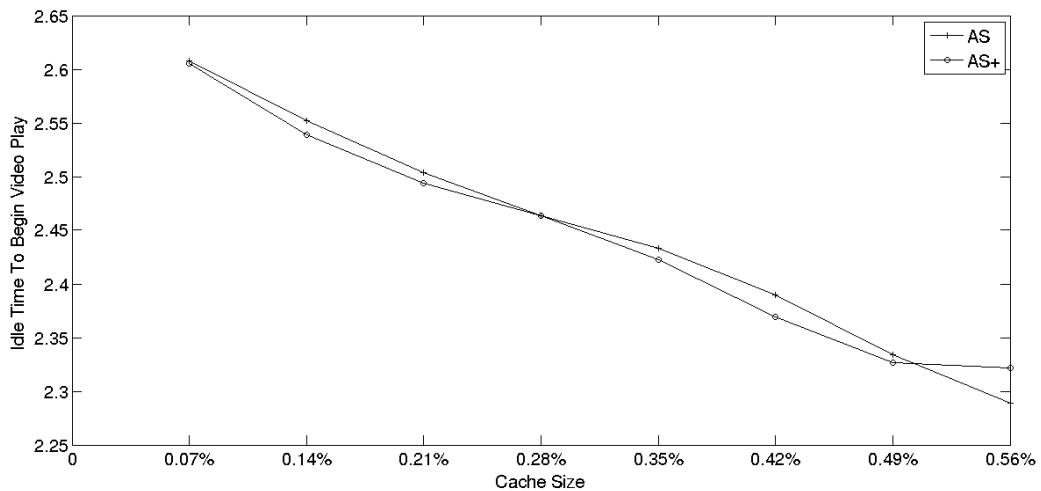
Στη περίπτωση της μη-προσαρμοσμένης ροής, οι ανενεργοί χρόνοι (μετρημένοι σε δευτερόλεπτα) μέχρι την εκκίνηση της αναπαραγωγής του βίντεο σε σχέση με το μέγεθος της κρυφής μνήμης των κόμβων του δικτύου καταγράφονται στο Σχήμα 16. Όπως είναι αναμενόμενο, οι χρόνοι αυτοί αυξάνονται με την ποιότητα και μειώνονται με την αύξηση του μεγέθους της κρυφής μνήμης. Εντούτοις, η βελτίωση είναι σχετικά μικρή (στην καλύτερη περίπτωση της μέγιστης ποιότητας υπάρχει βελτίωση της τάξης των 2 δευτερολέπτων).

Μένοντας στη μη-προσαρμοσμένη ροή βίντεο, οι χρόνοι παγώματος του βίντεο κατά τη διάρκεια της αναπαραγωγής του στην εφαρμογή του χρήστη απεικονίζονται στο Σχήμα 17. Όπως ξεκάθαρα φαίνεται και αναμενόταν, οι χρόνοι αυτοί μειώνονται και πάλι με την αύξηση της κρυφής μνήμης και αυξάνονται με την ποιότητα. Ένα αξιοσημείωτο γεγονός είναι η απότομη αύξηση των ανενεργών χρόνων πηγαινόντας στην ποιότητα Q_7 , η οποία ισοδυναμεί με 1800kbps. Αυτό οφείλεται στο ότι η γραμμή των 2Mbps πλέον στραγγίζει, με αποτέλεσμα την απότομη αύξηση των χρόνων. Το συμπέρασμα αυτό επιβεβαιώνεται και από το γεγονός ότι από αυτό το σημείο και πέρα, η αύξηση είναι και πάλι ομαλή, αφού έχει ξεπεραστεί το εν λόγω κατώφλι. Οι χρόνοι για τις χαμηλότερες ποιότητες είναι αμελητέοι και, συνεπώς, δε δίνονται στη γραφική παράσταση.

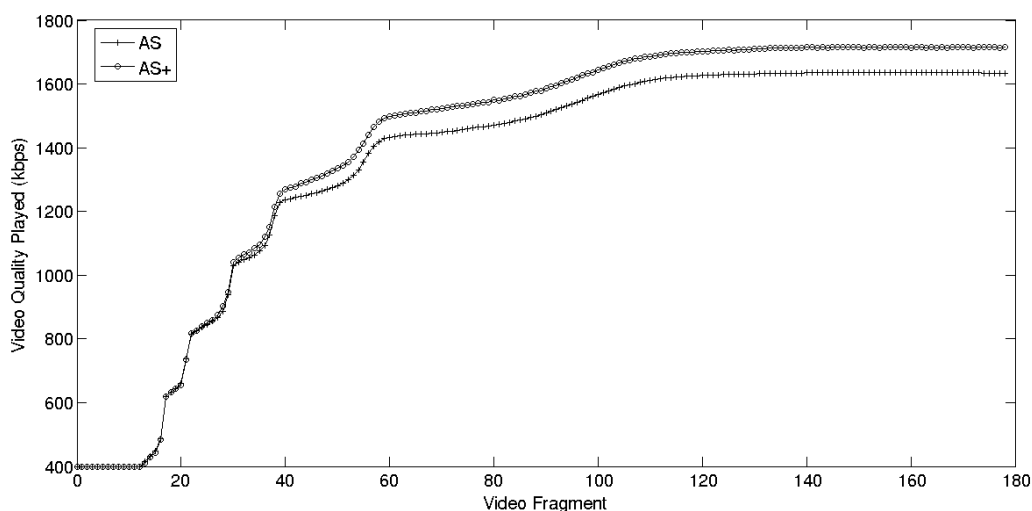
Στην προσαρμοσμένη ροή βίντεο με γραμμή 2Mbps και σε ό,τι αφορά το χρόνο εκκίνησης της αναπαραγωγής, τα πράγματα δεν είναι διαφορετικά σε σχέση με ό,τι παρατηρήθηκε και παραπάνω για τη μη-προσαρμοσμένη ροή. Συγκεκριμένα, όπως φαίνεται και



Σχήμα 17: Χρόνος Αναμονής κατά την Αναπαραγωγή στη Μη-Προσαρμοσμένη Ροή



Σχήμα 18: Χρόνος Αναμονής για Εκκίνηση Αναπαραγωγής στην Προσαρμοσμένη Ροή (2Mbps)



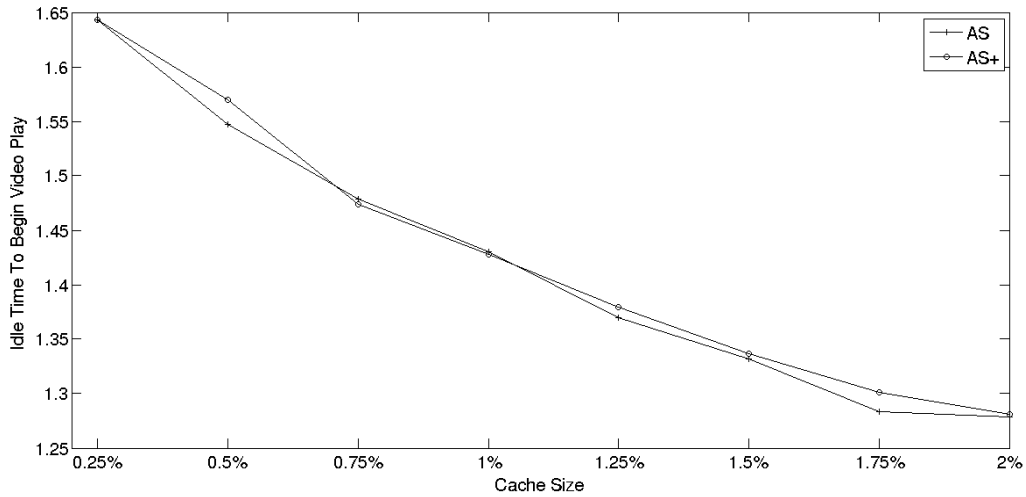
Σχήμα 19: Ποιότητα Αναπαραγωγής Βίντεο στην Προσαρμοσμένη Ροή (2Mbps)

στο Σχήμα 18, οι χρόνοι αναμονής είναι αντίστοιχοι με αυτούς της βασικής ποιότητας της μη-προσαρμοσμένης ροής, δεδομένου ότι και εδώ, η αναπαραγωγή ξεκινά από αυτήν την ποιότητα. Όπως είναι φυσικό, ο χρόνος αυτός μειώνεται με την αύξηση του μεγέθους της κρυφής μνήμης. Η μείωση δεν είναι αισθητή (της τάξης των εκατοστών του δευτερολέπτου), δεδομένου ότι στην αρχή δεν έχουν προλάβει να γεμίσουν οι κρυφές μνήμες.

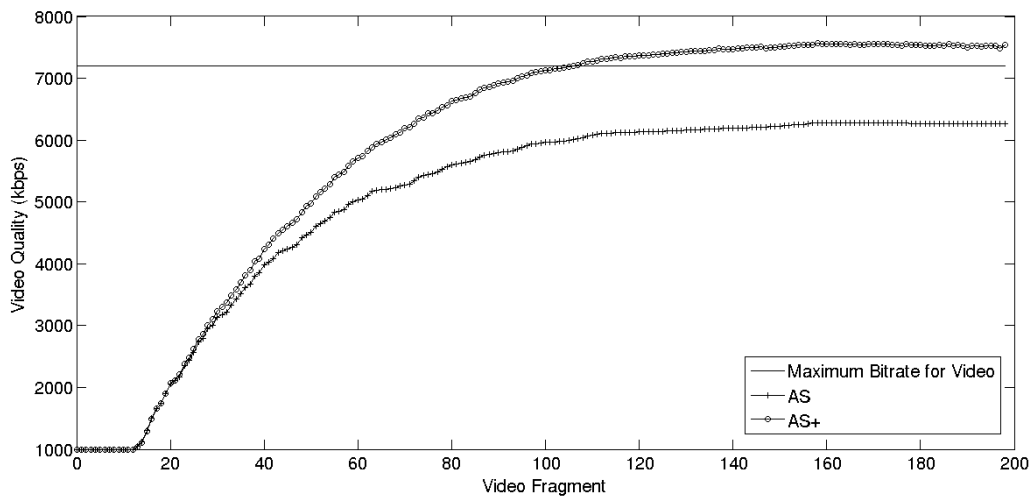
Η αντίστοιχη ποιότητα που απολαμβάνει ο χρήστης με την προσαρμοσμένη ροή βίντεο στην τυπική (AS) και την επεκτεταμένη (AS+) περίπτωση απεικονίζεται στο Σχήμα 19. Οι τιμές που χρησιμοποιήθηκαν αφορούν το μέγιστο μέγεθος κρυφής μνήμης. Όπως φαίνεται, η διαφορά στην επίδοση δεν είναι ιδιαίτερα μεγάλη, πράγμα που οφείλεται στο έτσι κι αλλιώς μικρό, συνολικά, μέγεθος της μνήμης που θεωρήθηκε. Ακόμα και έτσι, όμως, η επίδοση της προτεινόμενης επέκτασης είναι καλύτερη κατά περίπου 100kbps όταν συγκρίνει το πρωτόκολλο.

Οι παραπάνω μετρήσεις για την προσαρμοσμένη ροή και, συγκεκριμένα, για το χρόνο εκκίνησης αναπαραγωγής του βίντεο καθώς και την ποιότητα που απολαμβάνει ο τελικός χρήστης έγιναν και για δίκτυο με εύρος ζώνης της τάξης των 8Mbps. Έτσι, το Σχήμα 20 παρουσιάζει τα αντίστοιχα αποτελέσματα για την τυπική και την επεκτεταμένη περίπτωση. Οι χρόνοι παρουσιάζουν αντίστοιχη συμπεριφορά με τη μικρότερη γραμμή των 2Mbps με τη διαφορά ότι σε αυτήν την περίπτωση είναι μικρότεροι, δεδομένου ότι το μεγαλύτερο εύρος ζώνης επιφέρει και ταχύτερη μεταφορά του πρώτου τμήματος του βίντεο.

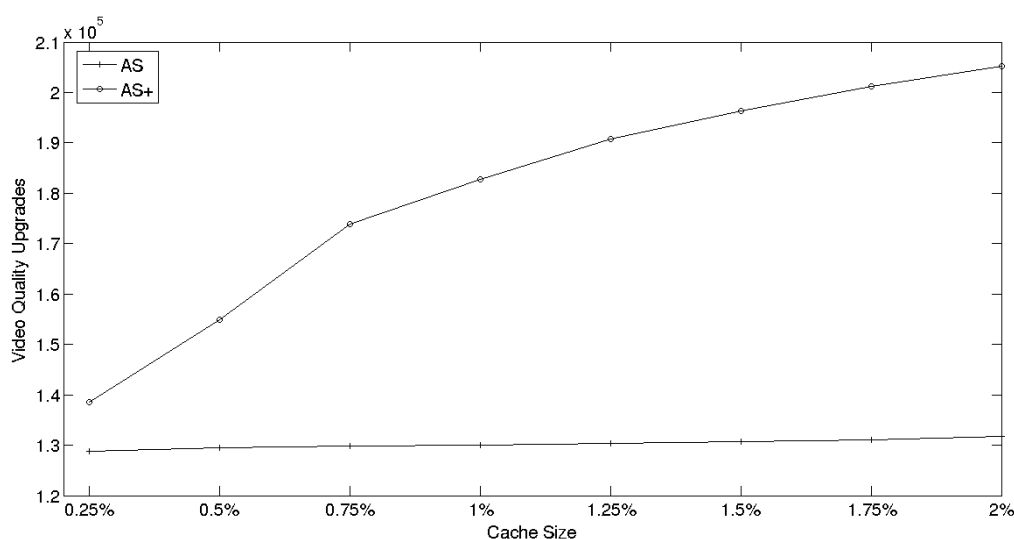
Ιδιαίτερο ενδιαφέρον, όμως, έχει η σύγκριση της ποιότητας στην τυπική και την επεκτεταμένη περίπτωση στα 8Mbps. Όπως φαίνεται στο Σχήμα 21, όπου έχει χρησιμοποιηθεί και κρυφή μνήμη μεγαλύτερου μεγέθους (2% του συνολικού περιεχομένου ελάχιστης ποιότητας), η προτεινόμενη επέκταση αποδίδει ποιότητα πάνω από 1200kbps μεγαλύτερη. Μάλιστα, αυτό που έχει εξαιρετική σημασία είναι ότι η απόδοση του πρωτοκόλλου στην



Σχήμα 20: Χρόνος Αναμονής για Εκκίνηση Αναπαραγωγής στην Προσαρμοσμένη Ροή (8Mbps)



Σχήμα 21: Ποιότητα Αναπαραγωγής Βίντεο στην Προσαρμοσμένη Ροή (8Mbps)

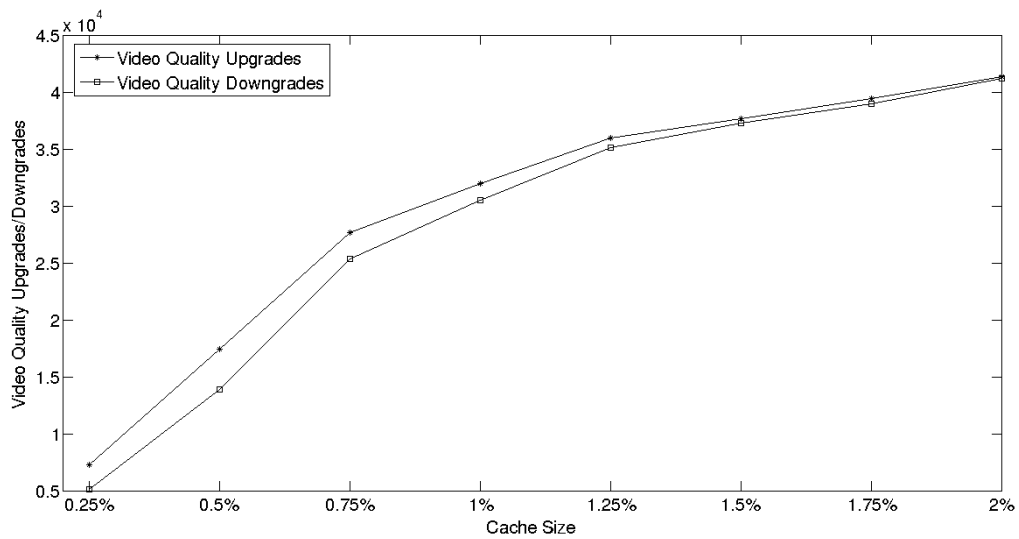


Σχήμα 22: Αναβαθμίσεις στην Ποιότητα Αναπαραγωγής Βίντεο (8Mbps)

πραγματικότητα φτάνει πολύ κοντά στα όρια της γραμμής, ειδικά με δεδομένο ότι υπάρχει κυμαινόμενη κίνηση μέχρι τα 800kbps για άλλες εφαρμογές πλην της ροής βίντεο, όπως φαίνεται στο σχήμα.

Η διαφορά στην επίδοση του πρωτοκόλλου από την τυπική στην επεκτεταμένη περίπτωση αντικατοπτρίζεται στο Σχήμα 22, το οποίο παρουσιάζει τις αναβαθμίσεις στην ποιότητα που γίνονται στις δύο αυτές περιπτώσεις, σε σχέση με το μέγεθος της κρυφής μνήμης. Όσο η τελευταία μεγαλώνει, τόσο μεγαλώνει και η διαφορά μεταξύ των δύο προσεγγίσεων. Αυτό οφείλεται στο ότι η προτεινόμενη επέκταση του πρωτοκόλλου κάνει καλύτερη διαχείριση της κρυφής μνήμης, αφού δεν επιτρέπει επικαλύψεις μεταξύ ίδιων τμημάτων του ίδιου βίντεο.

Τέλος, στο Σχήμα 23 παρουσιάζεται ο αριθμός των αυτόματων αναβαθμίσεων και υποβαθμίσεων που λαμβάνουν χώρα εξαιτίας της προτεινόμενης επέκτασης. Όσο μεγαλώνει το μέγεθος της κρυφής μνήμης, τόσο συγκλίνουν οι τιμές αυτών των δύο, με αποτέλεσμα να υπάρχει ουδέτερη, συνολικά, επίδραση στην τελική ποιότητα. Η αύξηση της ποιότητας που παρατηρήθηκε στο Σχήμα 21, οφείλεται στην καλύτερη διαχείριση της κρυφής μνήμης που δίνει τελικά τη δυνατότητα στο πρωτόκολλο να γεμίζει γρηγορότερα τους ενταμιευτές και, τελικά, να ανεβαίνει σε ποιότητα. Το Σχήμα 23 καθιστά σαφές ότι δεν είναι ότι εξυπηρετούνται περισσότερο υψηλότερες ποιότητες και η βελτίωση έχει να κάνει καθαρά με την καλύτερη κατανομή του περιεχομένου στις τοπικές μνήμες των κόμβων του δικτύου.



Σχήμα 23: Αναβαθμίσεις/Υποβαθμίσεις στην Ποιότητα Αναπαραγωγής Βίντεο (8Mbps)

Κεφάλαιο 3

Το Πρωτόκολλο Κατανεμημένης Ανακάλυψης Περιεχομένου CCPS

3.1 Εισαγωγή

Το πρόβλημα της δημοσίευσης/συνδρομής (publish/subscribe) αποτελεί μια ενεργή ερευνητική περιοχή των κατανεμημένων συστημάτων για αρκετά χρόνια τώρα και, ανάλογα με το πεδίο εφαρμογής, τα προτεινόμενα πρωτόκολλα μπορούν να κατηγοριοποιηθούν [16] ως εξής:

- Πρωτόκολλα βασιζόμενα σε κάποιο θέμα (topic-based): τα θέματα συνήθως αναπαρίστανται από κάποιο μοναδικό αναγνωριστικό που αντιπροσωπεύει μια ομάδα από ενδιαφερόμενους και διατίθεται στις περιπτώσεις ως λέξεις-κλειδιά. Σε αυτά τα συστήματα, κάθε συνδρομητής ενός συγκεκριμένου θέματος γίνεται αυτομάτως μέλος της αντίστοιχης ομάδας που συναπαρτίζεται από όλους τους συνδρομητές του εν λόγω θέματος.
- Πρωτόκολλα βασιζόμενα στο περιεχόμενο (content-based): οι συνδρομές εκφράζονται μέσω πληροφορίας, η οποία θα πρέπει να ανταποκρίνεται στο δημοσιευμένο περιεχόμενο, ακόμα και όχι πλήρως. Αυτή η οικογένεια πρωτοκόλλων θεωρείται πιο ευέλικτη, καθώς δεν απαιτείται κάποιο προκαθορισμένο σχήμα για θέματα, όπως στην προηγούμενη περίπτωση.
- Πρωτόκολλα βασιζόμενα στον τύπο (type-based): δε διαφέρουν ιδιαίτερα από την πρώτη οικογένεια. Η διαφορά έγκειται στο ότι εδώ δεν υπάρχει κάποιο εξωτερικά προσδιορισμένο, μοναδικό αναγνωριστικό. Αντιθέτως, κάθε συνδρομή εκφράζεται μέσα από κάποιον τύπο, ο οποίος εξάγεται ως συμπέρασμα από την ίδια τη συνδρομή.

Τα συστήματα που προκύπτουν από τις τρεις παραπάνω οικογένειες μπορούν, φυσικά, να συνδυαστούν κατά περίπτωση, προκειμένου να επιτευχθεί καλύτερη απόδοση.

Για παράδειγμα, οι συμμετέχοντες μπορεί να εγγράφονται αρχικά σε κάποιο θέμα και στη συνέχεια, αφού έχουν (εικονικά) τοποθετηθεί σε κάποια ευρύτερη ομάδα να πραγματοποιούν συνδρομές με βάση το περιεχόμενο, προκειμένου να εκμεταλλευτούν την ευελιξία που παρέχουν τα εν λόγω πρωτόκολλα. Το πρωτόκολλο που παρουσιάζεται στην παρούσα διατριβή επικεντρώνεται στο περιεχόμενο (content-based), με τις συνδρομές να φέρουν ένα ερώτημα εκπεφρασμένο με λέξεις-κλειδιά (keyword-based query), ώστε να υποστηρίζεται και μερικό ταιρίασμα (partial matching) με το δημοσιευμένο περιεχόμενο.

Μολονότι υπάρχουν διάφοροι τρόποι να υποστηριχθούν μηχανισμοί για συνδρομές με βάση το περιεχόμενο, γενικά προτείνεται η χρήση επιδημικών πρωτοκόλλων, τα οποία μπορούν να υποστηρίξουν ακόμα και μερικό ταιρίασμα των ερωτημάτων που φέρουν οι συνδρομές, καθώς ο μηχανισμός ταιριάσματος υλοποιείται σε κάθε κάτοχο δημοσιευμένου περιεχομένου. Αυτή η περίπτωση είναι ιδιαίτερα επιθυμητή όταν ο τελευταίος υλοποιεί προχωρημένους μηχανισμούς ταιριάσματος, π.χ. βασισμένους σε σημασιολογικά μεταδεδομένα, οπότε θέλει να αναλάβει ο ίδιος να εκτελέσει το ερώτημα στη δημοσιευμένη πληροφορία του.

Ένα τυπικό σύστημα δημοσίευσης/συνδρομής για ανακάλυψη περιεχομένου θεωρεί ότι κάποιος συνδρομητής εκφράζει το ενδιαφέρον του για κάποια πληροφορία και λαμβάνει ενημερώσεις μόλις η αντίστοιχη πληροφορία γίνει διαθέσιμη. Στα πλαίσια της διατριβής, αυτή η προσέγγιση επεκτάθηκε ώστε οι ενημερώσεις που φτάνουν στο συνδρομητή να αφορούν και πληροφορία που έχει ήδη δημοσιευτεί, ακολουθώντας λογική πρωτοκόλλου αναζήτησης.

Ένα κοινό πρόβλημα που παρουσιάζουν τα πρωτόκολλα αυτής της οικογένειας σε κατανεμημένα συστήματα και, ειδικά, σε ομότιμα δίκτυα είναι αυτό της κλιμάκωσης. Σε αντίθεση με τις προσεγγίσεις βασιζόμενες σε κάποιο θέμα ή τύπο, όπου η οργάνωση των συμμετεχόντων σε ομάδες μπορεί να γίνει με σχετικά πιο τετριμμένους τρόπους, στις βασιζόμενες σε περιεχόμενο η επίτευξη της κλιμάκωσης μπορεί να γίνει αρκετά δύσκολη. Από την άλλη όμως, οι τελευταίες είναι πιο ευέλικτες και αποδοτικές από την πλευρά του τελικού χρήστη, καθώς μπορούν να αποτυπώσουν με μεγαλύτερη ακρίβεια τα ενδιαφέροντά του, ενώ δύνανται να υποστηρίξουν και άλλους προχωρημένους μηχανισμούς αναζήτησης, όπως η σημασιολογική.

3.2 Θεμελιώδεις Σχεδιαστικές Αρχές του CCPS

Το πρωτόκολλο Caching Content-based Publish/Subscribe (CCPS) [40] που παρουσιάζεται στα πλαίσια της τρέχουσας διατριβής έχει σχεδιαστεί με τέτοιο τρόπο ώστε να αντιμετωπίζει τα προβλήματα κλιμάκωσης και εξισορρόπησης φόρτου στα συστήματα δημοσίευσης/συνδρομής που βασίζονται στο περιεχόμενο, εκμεταλλευόμενο την επαναληπτικότητα που παρουσιάζουν τα ερωτήματα των χρηστών. Αυτό επιτυγχάνεται μέσω της χρή-

σης κρυφής μνήμης (cache), όπου αποθηκεύονται πληροφορίες σχετικές με τους συνδρομητές που μοιράζονται κοινά ενδιαφέροντα, τα οποία εκφράζονται μέσα από τα ερωτήματα συνδρομής. Η κρυφή μνήμη υλοποιείται ως ένας κατανεμημένος πίνακας κατακερματισμού (Distributed Hash Table - DHT), ώστε να κλιμακώνει, ενώ, όπως αποδεικνύεται στη συνέχεια του κεφαλαίου, η απόδοσή της αυξάνεται με την πιθανότητα επανάληψης ενός ερωτήματος. Η περιγραφή του πρωτοκόλλου συνοδεύεται τόσο από ένα θεωρητικό μοντέλο, το οποίο βασίζεται στις παρατηρήσεις επαναληπτικότητας των ερωτημάτων που έχουν γίνει στο [41], όσο και από προσομοιώσεις, οι οποίες επιβεβαιώνουν το μοντέλο αποδεικνύοντας ότι η βελτίωση στην επιβάρυνση του συστήματος από μηνύματα ανακάλυψης ισούται με την πιθανότητα να υπάρχει επιτυχές χτύπημα στην κρυφή μνήμη και φτάνει μέχρι την πιθανότητα επανάληψης ενός ερωτήματος.

Μια ακόμα συνεισφορά του CCPS αφορά το πρόβλημα κατανομής φόρτου, προτείνοντας μια αρχιτεκτονική όπου οι συνδρομητές του ίδιου ερωτήματος μοιράζονται μεταξύ τους το αντίστοιχο περιεχόμενο. Ο διαμοιρασμός αυτός πραγματοποιείται με χρήση ομότιμων πρωτοκόλλων, όπως το BitTorrent. Κατά αυτόν τον τρόπο, οι οντότητες του συστήματος που δημοσιεύουν περιεχόμενο δε χρειάζεται να στέλνουν οι ίδιοι κάθε φορά το περιεχόμενο αυτό, από τη στιγμή που το διαμοιράζονται οι συνδρομητές. Επιπλέον, οι τελευταίοι κατανέμουν με δίκαιο τρόπο (ανάλογα φυσικά και με το πρωτόκολλο που χρησιμοποιούν για το σκοπό αυτό) το φόρτο διαμοιρασμού. Στο υπόλοιπο αυτής της διατριβής θα θεωρηθεί ότι για το διαμοιρασμό χρησιμοποιείται το BitTorrent, χωρίς αυτό να σημαίνει ότι το CCPS έχει αποκλειστική εξάρτηση από το BitTorrent.

Η χρήση της κρυφής μνήμης έχει ένα ακόμα πλεονέκτημα ως προς την απόδοση του συστήματος. Αρχικά, η αποστολή των ερωτημάτων (συνδρομών) στις οντότητες που δημοσιεύουν το περιεχόμενο γίνεται με τη χρήση επιδημικών πρωτοκόλλων, όπως προτείνεται στη βιβλιογραφία [42] [43] [44]. Εντούτοις, εξαιτίας της πιθανοτικής φύσης των επιδημικών πρωτοκόλλων, μπορεί ένα ερώτημα να φτάσει σε διάφορους εξυπηρετητές ίδιου περιεχομένου, με αποτέλεσμα η ίδια απάντηση να σταλεί στον ίδιο συνδρομητή από διαφορετικές οντότητες του συστήματος, το οποίο οδηγεί σε σπατάλη εύρους ζώνης και ισχύος. Στο CCPS κάθε συνδρομητής διατηρεί μόνο ένα αντίγραφο του δημοσιευμένου περιεχομένου, οπότε ανεξαρτήτως του πόσες φορές έχει δημοσιευτεί κάθε επόμενος συνδρομητής θα λάβει μόνο ένα αντίγραφο του. Επιπλέον, από τη στιγμή που αποφεύγονται επιδημικοί γύροι εξάπλωσης ερωτημάτων, οι κάτοχοι του περιεχομένου ούτε λαμβάνουν ούτε απαντούν σε ερωτήματα πάνω από μια φορά (τουλάχιστο στην ιδανική περίπτωση).

3.3 Τρέχουσες Τάσεις στην Ευρύτερη Ερευνητική Περιοχή

Διάφορα πρωτόκολλα που επικεντρώνονται σε διαφορετικές πτυχές του προβλήματος δημοσίευσης/συνδρομής έχουν προταθεί στη βιβλιογραφία, με την πλειονότητά τους να θεωρεί ότι οι συνδρομές θα πρέπει να αφορούν μεταγενέστερες αυτών δημοσιεύσεις.

Στην παρούσα διατριβή επεκτείνεται αυτή η υπόθεση και σε παρελθοντικές δημοσιεύσεις καθιστώντας το σύστημα εφάμιλλο ενός γενικότερου συστήματος αναζήτησης σε κατανεμημένα συστήματα. Επιπλέον, οι οντότητες του συστήματος που δημοσιεύουν περιεχόμενο είναι οργανωμένοι σε ένα ομότιμο υπερκείμενο δίκτυο, ενώ οι συνδρομητές μπορούν να πραγματοποιούν ερωτήματα με βάση λέξεις-κλειδιά, υποστηρίζοντας μερικό ταίριασμα πληροφορίας.

Συστήματα βασισμένα σε κατανεμημένους πίνακες κατακερματισμού (DHT), όπως το CAN [45], το Pastry [46] το Chord [18] και το Kademia [47] αποτελούν μια συνηθισμένη λύση στη βιβλιογραφία για προβλήματα δημοσίευσης/συνδρομής. Ωστόσο, αυτά τα πρωτόκολλα ακολουθούν μια προσέγγιση που βασίζεται σε δρομολόγηση με βάση κάποιο κλειδί και, συνεπώς, αδυνατούν να υποστηρίξουν ερωτήματα με λέξεις-κλειδιά καθώς και μερικό ταίριασμα, τα οποία αποτελούν τυπική απαίτηση συστημάτων δημοσίευσης/συνδρομής βασισμένων στο περιεχόμενο [48]. Από την άλλη μεριά, η δρομολόγηση των ερωτημάτων στους κατάλληλους κόμβους του συστήματος, ώστε να υποστηριχθεί η συνδρομή με βάση λέξεις-κλειδιά μπορεί να πραγματοποιηθεί είτε με αλγορίθμους πλημμύρας είτε με τυχαίους περιπάτους. Δεδομένου ότι οι τελευταίοι δεν έχουν καλά ποσοστά επιτυχίας [48], ένας συνδυασμός των δύο θα χρησιμοποιηθεί.

Τα ερωτήματα που βασίζονται σε λέξεις-κλειδιά συνήθως μεταδίδονται σε ένα υπερκείμενο ομότιμο δίκτυο με χρήση πιθανοτικών αλγορίθμων και, συγκεκριμένα, επιδημικών [49], οι οποίοι θεωρούνται κλιμακώσιμοι και ανεκτικοί σε βλάβες και χαμηλή συνδεσιμότητα [50], ενώ επιτρέπουν σε κάθε κόμβο που λαμβάνει ένα ερώτημα να πραγματοποιεί μόνος του το ταίριασμα στο δημοσιευμένο περιεχόμενο. Μια συνηθισμένη παραλλαγή των επιδημικών αλγορίθμων αφορά τον ιεραρχικό σχηματισμό των κόμβων [44], που οδηγεί σε καλύτερη απόδοση του συστήματος, ενώ παράλληλα επιτρέπει και την εισαγωγή επιπλέον τεχνολογιών που υπαγορεύουν αποδοτικές ιεραρχίες με βάση το εκάστοτε πρόβλημα.

Επιδημικοί αλγόριθμοι έχουν χρησιμοποιηθεί στα πλαίσια διαφόρων πρωτοκόλλων, τα οποία επικεντρώνονται κυρίως σε τρόπους να περιορίσουν τον αριθμό των κόμβων υποψηφίων για την παραλαβή ενός επιδημικού μηνύματος. Ένα χαρακτηριστικό παράδειγμα αποτελεί το Associative Search [51], όπου δημιουργούνται συνδέσεις μεταξύ κόμβων ανάλογα με την ομοιότητα των ενεργειών των αντίστοιχων χρηστών. Το Foreseer [52] θεωρεί ένα υπερκείμενο δίκτυο δύο επιπέδων, όπου το πρώτο επίπεδο δημιουργείται με βάση την εγγύτητα των κόμβων σε επίπεδο δικτύου και το δεύτερο με βάση τις απαντήσεις στα ερωτήματα των συνδρομητών. Στο SpiderCast [53], οι συγγραφείς επιχειρούν μια πιο γενικευμένη προσέγγιση στο πρόβλημα της δημοσίευσης/συνδρομής με χρήση επιδημικών αλγορίθμων και πραγματοποιούν εκτιμήσεις της απόδοσης του συστήματος υπό διάφορες συνθήκες. Το Vitis [54] βασίζεται στο T-MAN [55], το οποίο είναι ένα επιδημικό πρωτόκολλο για τη δημιουργία γειτονιών με βάση θέματα, και χρησιμοποιεί αυτές τις γειτονιές για να στείλει τα μηνύματα συνδρομών. Το StAN [56] επικεντρώνεται στη δημιουργία θεματικών συστάδων, σύμφωνα με τη φυσική τους τοποθεσία, ώστε να ελαχιστοποιείται ο αριθμός

των ενδιάμεσων δρομολογητών δικτύων στην επικοινωνία μεταξύ των αντίστοιχων κόμβων.

Δεδομένου ότι η κλιμάκωση είναι το βασικό πρόβλημα σε ένα επιδημικό σύστημα δημοσίευσης/συνδρομής, οι περισσότερες εργασίες στη βιβλιογραφία επιχειρούν να αναγνωρίσουν πιθανές σχέσεις μεταξύ των κόμβων, ώστε να μειώσουν το χώρο αναζήτησης και, συνεπώς, τον αριθμό των μηνυμάτων ανακάλυψης που πρέπει να αποσταλούν (βλ. [57] [58] [59] [60]). Μολονότι το CCPS έχει βασιστεί στην ίδια φιλοσοφία, η επαναληπτικότητα των ερωτημάτων μπορεί να θεωρηθεί τόσο ισχυρή όσο και γενική βάση για να σχεδιαστεί ένα πρωτόκολλο δημοσίευσης/συνδρομής. Επιπλέον, το προτεινόμενο πρωτόκολλο μπορεί να συνυπάρξει με άλλα ήδη υπάρχοντα, καθώς δεν υπάρχουν επικαλύψεις στη μεθοδολογία: τα τυπικά επιδημικά πρωτόκολλα περιορίζουν το χώρο αναζήτησης, ενώ το CCPS επιστρατεύει την κρυφή μνήμη για τη βελτίωση της απόδοσης.

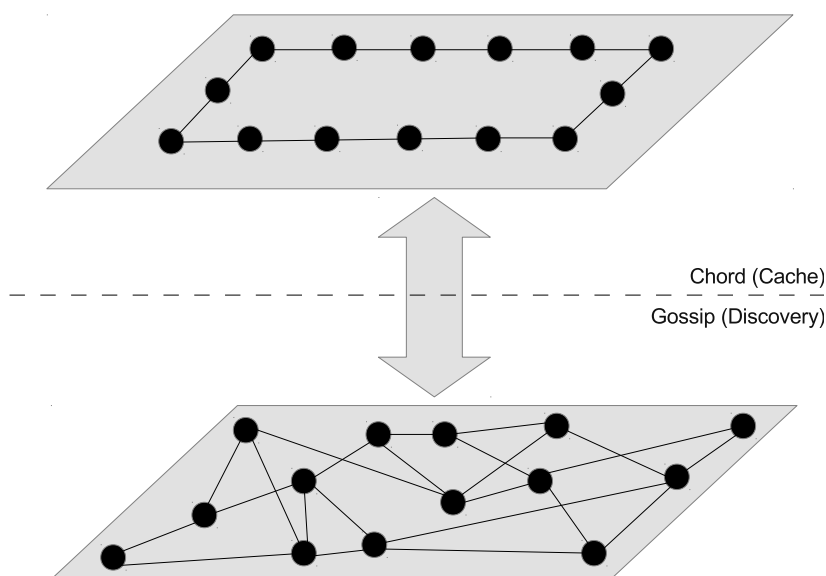
3.4 Περιγραφή Πρωτοκόλλου CCPS

3.4.1 Ανακάλυψη Περιεχομένου

Το προτεινόμενο πρωτόκολλο θεωρεί ότι οι κόμβοι που δημοσιεύουν περιεχόμενο είναι οργανωμένοι σε ένα υπερκείμενο δίκτυο και λαμβάνουν τις συνδρομές των χρηστών, τις ελέγχουν αν ταιριάζουν με την πληροφορία που έχουν δημοσιεύσει και στέλνουν το αντίστοιχο περιεχόμενο στους ενδιαφερόμενους. Οι συνδρομές φέρουν ερωτήματα, τα οποία είναι εκπεφρασμένα με λέξεις-κλειδιά και όχι με αναγνωριστικά, όπως συμβαίνει στα συστήματα βασισμένα σε κλειδιά (π.χ. DHT). Κατά αυτόν τον τρόπο, το CCPS δύναται να υποστηρίξει πολύπλοκα ερωτήματα καθώς και μερικό ταίριασμά τους, ενώ προσφέρεται η δυνατότητα στους χρήστες και για σημασιολογικά ερωτήματα. Επιπλέον, οι συνδρομές αφορούν τόσο υπάρχουσες όσο και μελλοντικές δημοσιεύσεις.

Μια ενδιαφέρουσα παρατήρηση για τις αναζητήσεις, που παρουσιάζεται στο [41], είναι ότι τα ερωτήματα αναζήτησης έχουν την τάση να επαναλαμβάνονται. Επιπλέον, η επαναληπτικότητά τους μεγαλώνει όσο το θέμα που αφορούν γίνεται πιο συγκεκριμένο. Λαμβάνοντας υπόψη ότι η μετάδοση του ερωτήματος στους κόμβους που δημοσιεύουν το περιεχόμενο εισάγει τα προβλήματα κλιμάκωσης στα συστήματα δημοσίευσης/συνδρομής βασισμένα στο περιεχόμενο, αυτή η παρατήρηση αποτέλεσε το κίνητρο για τη δημιουργία μιας κρυφής μνήμης για τα ερωτήματα που έχουν πραγματοποιηθεί και είναι πιθανό να επαναληφθούν και στο κοντινό μέλλον. Η κρυφή μνήμη περιλαμβάνει πληροφορία αντιστοίχισης ανάμεσα στα ερωτήματα και άλλους συνδρομητές που τα έχουν ήδη πραγματοποιήσει. Καθώς η μνήμη θα περιέχει ερωτήματα που έχουν ήδη πραγματοποιηθεί, αυτά μπορούν να αναπαρασταθούν με ένα μοναδικό αναγνωριστικό, καθιστώντας τις προσεγγίσεις DHT ως τις πιο ελκυστικές για αυτό το σκοπό.

Όταν κάποιος συνδρομητής επιθυμεί να πραγματοποιήσει ένα ερώτημα, τότε το



Σχήμα 24: Συνδρομή στο Υπερκείμενο Δίκτυο Δύο Επιπέδων Δημοσιευόντων Κόμβων

CCPS κάνει πρώτα μια αναζήτηση στην κρυφή μνήμη, η οποία είναι υλοποιημένη με DHT απαρτιζόμενο από τους κόμβους που δημοσιεύουν περιεχόμενο. Αν δε βρεθεί περιεχόμενο στην κρυφή μνήμη, τότε το ερώτημα στέλνεται με επιδημικό αλγόριθμο στους αντίστοιχους κόμβους. Το υπερκείμενο δίκτυο των τελευταίων μπορεί να οργανωθεί με διάφορους τρόπους προκειμένου να βελτιωθεί η απόδοση του συστήματος (κάποιοι τρόποι παρουσιάζονται στην Ενότητα 3.3), χωρίς αυτό να επηρεάζει τη λειτουργία του πρωτοκόλλου. Αφού έχει μαζέψει όλα τα αποτελέσματα, ο συνδρομητής αποθηκεύει τη σχετική πληροφορία στην κρυφή μνήμη. Έτσι, όταν στο μέλλον κάποιος άλλος συνδρομητής πραγματοποιήσει το ίδιο ερώτημα, θα μοιραστεί τα αποτελέσματα με τον ήδη υπάρχοντα, ενώ προστίθεται και αυτός στη σχετική λίστα συνδρομητών που διαθέτουν αυτό το περιεχόμενο.

Το Σχήμα 24 παρουσιάζει τη διεπίπεδη οργάνωση των κόμβων που δημοσιεύουν περιεχόμενο, ώστε να υποστηριχθεί η περιγραφείσα λειτουργία του πρωτοκόλλου. Το πρώτο επίπεδο απεικονίζει την κατανεμημένη κρυφή μνήμη: όταν κάποιος συνδρομητής κάνει μια αίτηση που μεταφέρει ένα ερώτημα, αυτό αναζητείται αρχικά στην κρυφή μνήμη. Σε περίπτωση που δεν υπάρχει αποτέλεσμα, ο συνδρομητής απευθύνεται στο δεύτερο επίπεδο και αναζητά περιεχόμενο που να ικανοποιεί το ερώτημα σε όλους τους κατόχους περιεχομένου. Όταν συλλέξει όλη τη σχετική πληροφορία, ο συνδρομητής αφαιρεί όλα τυχών επικαλυπτόμενα αποτελέσματα (επικάλυψη μπορεί να υπάρχει επειδή π.χ. μια ταινία δημοσιεύεται από περισσότερους του ενός κόμβους), δημιουργεί ένα αρχείο μετα-πληροφορίας, το οποίο φέρει ένα μοναδικό αναγνωριστικό που ανταποκρίνεται στο ερώτημα (π.χ. το αποτέλεσμα της συνάρτησης σύνοψης που χρησιμοποιεί το DHT) και το αποθηκεύει στην κρυφή μνήμη. Έτσι, όταν στο μέλλον κάποιος άλλος συνδρομητής πραγματοποιήσει το ίδιο

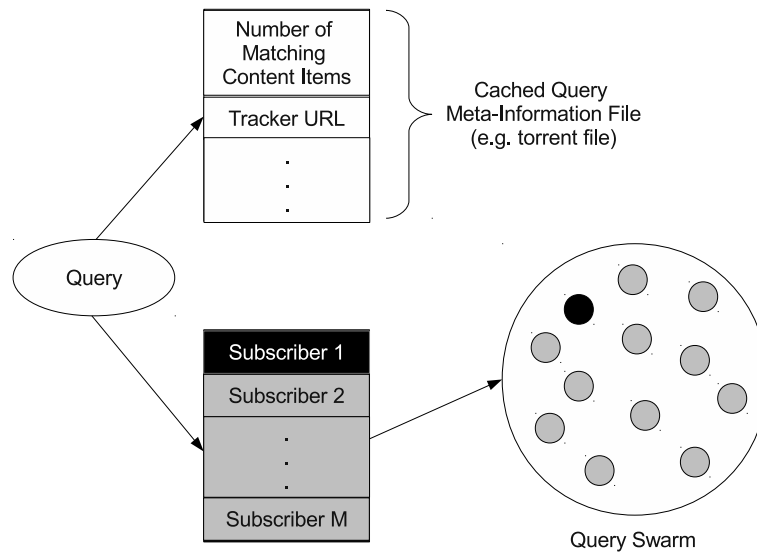
ερώτημα, παίρνει αυτό το αρχείο από την κρυφή μνήμη και αρχίζει να λαμβάνει τα αποτελέσματα από τους υπόλοιπους συνδρομητές που έχουν πραγματοποιήσει το ίδιο ερώτημα στο παρελθόν.

Η περίπτωση που μελετάται στο πλαίσιο της παρούσας διατριβής βασίζεται στις παρατηρήσεις που έχουν γίνει στο [41], καθώς και σε στατιστικά στοιχεία του Google trends [61] και θεωρεί ότι κάθε ερώτημα επαναλαμβάνεται στην ίδια (αρχική) μορφή. Εναλλακτικά, συστήματα σημασιολογικών τεχνολογιών δύνανται να εφαρμοστούν, όπου τα ερωτήματα μπορούν να μετασχηματιστούν σε μια καθολική μορφή, με βάση κάποια οντολογία ανώτερης ιεραρχίας.

Μολονότι τα ερωτήματα θα μπορούσαν να χρησιμοποιηθούν εξ αρχής ώστε να ακολουθηθεί μια καθαρή προσέγγιση βασισμένη σε DHT, κάτι τέτοιο δεν είναι εφικτό καθώς δεν είναι γνωστά πριν εφαρμοστούν. Χρησιμοποιώντας την κρυφή μνήμη και βασισμένο στην επαναληπτικότητα των ερωτημάτων, το σύστημα είναι ικανό να προσαρμόζεται, ώστε να εξυπηρετεί τα συχνά ερωτήματα με πιο αποδοτικό τρόπο. Έστω, για παράδειγμα, μια δημοσίευση για μια ρομαντική ταινία που τη σκηνοθετεί ο John Doe, η οποία έχει γυριστεί στη Σαντορίνη. Διάφορα ερωτήματα θα μπορούσαν να γίνουν και να έχουν ως αποτέλεσμα αυτήν ταινία, όπως “romantic movie”, “movie directed by John Doe” και “movie shot in Santorini”. Επιπλέον, ερωτήματα του τύπου “movie shot in Greek island”, “movie shot in island”, “movie shot in island of the Aegean” κλπ επίσης μπορούν να ταιριάζουν στο εν λόγω δημοσιευμένο περιεχόμενο. Η εξαγωγή όλων των πιθανών ερωτημάτων εξ αρχής, ώστε να χρησιμοποιηθεί μόνο ένα σύστημα βασισμένο σε κλειδιά (DHT) δε θα ήταν αποδοτική. Από την άλλη μεριά, θεωρώντας ότι το ερώτημα “romantic movie” πραγματοποιείται από πολλούς χρήστες, τότε μόνο αυτό θα χρησιμοποιηθεί για το DHT και θα αντιστοιχηθεί και με την εν λόγω δημοσίευση.

Το CCPS υλοποιεί την κρυφή μνήμη με το πρωτόκολλο Chord [18]. Το Chord δημιουργεί ένα υπερκείμενο δίκτυο, κάθε κόμβος του οποίου φέρει ένα μοναδικό αναγνωριστικό, το οποίο συνήθως είναι το αποτέλεσμα της συνάρτησης κατακερματισμού που χρησιμοποιείται πάνω στο όνομα της υπηρεσίας που παρέχει ο κόμβος αυτός για πρόσβαση στην πληροφορία που δημοσιεύει. Επιπλέον, κάθε κόμβος διατηρεί ένα πίνακα δρομολόγησης που ονομάζεται *finger* και έχει $\log N$ εγγραφές, όπου N ο συνολικός αριθμός των κόμβων που συμμετέχουν στο υπερκείμενο δίκτυο. Όταν ένα τμήμα περιεχομένου (π.χ. αρχείο) εισάγεται στο Chord, το αναγνωριστικό του εφαρμόζεται στη συνάρτηση κατακερματισμού και γίνεται το κλειδί που χρησιμοποιείται για την αναζήτησή του. Στη συνέχεια, με βάση το *finger*, το πρωτόκολλο αναζητά τον κόμβο που έχει αναγνωριστικό ίσο ή αμέσως μεγαλύτερο από αυτό του περιεχομένου. Η αναζήτηση αυτή γίνεται σε $O(\log N)$ βήματα. Αυτός ο κόμβος λέγεται *successor* και είναι υπεύθυνος να αποθηκεύει πληροφορία με αναγνωριστικά ίσα ή μεγαλύτερα του δικού του. Αντίστοιχη διαδικασία ακολουθείται και όταν αναζητείται περιεχόμενο που φέρει κάποιο αναγνωριστικό.

Στην ουσία, η προτεινόμενη προσέγγιση θεωρεί ότι ένα ερώτημα συνδρομής αντι-



Σχήμα 25: Διαμοιρασμός Γνώσης Μεταξύ των Συνδρομητών

προσωπεύει ένα κομμάτι γνώσης, με όλα τα τμήματα περιεχομένου που απαντούν στο ερώτημα να αποτελούν κομμάτια της γνώσης αυτής. Συνεπώς, όλοι οι κόμβοι συνδρομητές που κάνουν το ίδιο ερώτημα μπορούν να θεωρηθούν ως κάτοχοι της αντίστοιχης γνώσης και τη διαμοιράζονται μεταξύ τους.

3.4.2 Διαμοιρασμός Γνώσης

Ένα κομμάτι γνώσης αναγνωρίζεται από ένα ερώτημα και αποτελείται από όλα τα κομμάτια πληροφορίας που απαντούν σε αυτό το ερώτημα. Αυτά τα κομμάτια πληροφορίας μπορούν να θεωρηθούν (ή ακόμα και να είναι στην πράξη) μεμονωμένα αρχεία, τα οποία διαμοιράζονται ανάμεσα στους συνδρομητές του ίδιου ερωτήματος. Συνεπώς, οι δημοσιεύσεις δεν είναι απαραίτητο να προέρχονται από τους ίδιους τους κόμβους που τις κάνουν. Αρκεί οι συνδρομητές να τις μοιράζονται κάνοντας χρήση κάποιου ομότιμου πρωτοκόλλου, όπως είναι το BitTorrent.

Αυτή η προσέγγιση έχει άμεση θετική συνέπεια στην εξισορρόπηση φόρτου του συστήματος. Σε ένα τυπικό σύστημα δημοσίευσης/συνδρομής ή αναζήτησης, ο κόμβος που δημοσιεύει το περιεχόμενο είναι επιφορτισμένος να στέλνει και τα αντίστοιχα αποτελέσματα στους συνδρομητές. Το CCPS, όχι μόνο αποφορτίζει τους δημοσιεύοντες κόμβους από αυτή τη διαδικασία, αλλά επιπλέον κατανέμει και το φόρτο της τελικής παράδοσης του περιεχομένου στους υπάρχοντες συνδρομητές, που το έχουν ήδη λάβει. Μάλιστα, όσο πιο δημοφιλές είναι ένα ερώτημα, τόσο περισσότεροι είναι και οι συνδρομητές και, άρα, πιο καλή η κατανομή φόρτου.

Το Σχήμα 25 παρουσιάζει τον τρόπο με τον οποίο το CCPS υποστηρίζει το διαμοιρασμό γνώσης. Για κάθε ερώτημα δημιουργείται ένα αρχείο μετα-πληροφορίας, το οποίο υποδεικνύει τον αριθμό των στοιχείων περιεχομένου (κάθε δημοσίευση θεωρείται ότι αφορά ένα στοιχείο περιεχομένου) που απαντούν το εν λόγω ερώτημα και, αφαιρετικά, μπορεί να θεωρηθεί ανάλογο ενός αρχείου torrent. Το αρχείο αυτό μπορεί να περιλαμβάνει και επιπλέον πληροφορίες, όπως το URL του tracker³, κανόνες ελέγχου πρόσβασης στα προστατευόμενα τμήματα του περιεχομένου κλπ. Το αρχείο μετα-πληροφορίας δημιουργείται από κάποιο συνδρομητή, όταν αποτύχει να βρει αντίστοιχο στην κρυφή μνήμη. Συγκεκριμένα, μετά από μια τέτοια αποτυχία, ο συνδρομητής κάνει χρήση ενός επιδημικού πρωτοκόλλου αναζητώντας από κάθε ένα δημοσιεύων κόμβο περιεχόμενο που απαντά στο ερώτημά του. Στη συνέχεια, δημιουργεί το αρχείο μετα-πληροφορίας και το αποθηκεύει στην κατανεμημένη κρυφή μνήμη. Στη σημειολογία του BitTorrent, ο κόμβος αυτός θεωρείται *seeder* αυτής της γνώσης, αφού διαθέτει όλα τα επιμέρους κομμάτια που τη συναποτελούν.

Όταν ένας νέος συνδρομητής πραγματοποιεί το ίδιο ερώτημα, τότε λαμβάνει το αντίστοιχο αρχείο μετα-πληροφορίας από την κρυφή μνήμη και εξάγει τον αριθμό των στοιχείων περιεχομένου που απαντούν στο ερώτημα και οποιαδήποτε άλλη πληροφορία μπορεί να χρειάζεται (π.χ. το URL του tracker). Σε αυτό το σημείο είναι έτοιμος να ξεκινήσει να κατεβάζει τα στοιχεία περιεχομένου από τους υπόλοιπους συνδρομητές του ίδιου ερωτήματος. Έτσι, όπως προαναφέρθηκε, επιτυγχάνεται κατανομή του φόρτου αποστολής της γνώσης σε όλους τους συνδρομητές.

Ένα ζήτημα που χρήζει ιδιαίτερης προσοχής σε αυτό το σημείο είναι ότι τα πρωτόκολλα δημοσίευσης/συνδρομής είναι εγγενώς ασύγχρονα, πράγμα που σημαίνει ότι μια συνδρομή μπορεί να έχει αποτελέσματα και αφού πραγματοποιηθεί για πρώτη φορά το ερώτημα από το συνδρομητή. Στην περίπτωση που νέα στοιχεία περιεχομένου καταστούν διαθέσιμα, το αρχείο μετα-πληροφορίας στην κρυφή μνήμη θα πρέπει να ενημερωθεί. Για το σκοπό αυτό, ο κόμβος που δημοσιεύει το περιεχόμενο, ο οποίος γνωρίζει κάποιους συνδρομητές (αυτούς που δεν έχουν εξυπηρετηθεί από την κρυφή μνήμη), θα επιχειρήσει να ειδοποιήσει σχετικά τους συνδρομητές που γνωρίζει, ενώ παράλληλα θα ενημερώσει και τον tracker με το νέο αρχείο μετα-πληροφορίας. Οι συνδρομητές που θα λάβουν το νέο αυτό στοιχείο περιεχομένου θα συγχρονιστούν ξανά με τον tracker, ώστε να πάρουν τη λίστα με τους υπόλοιπους συνδρομητές και θα τους αποστείλουν ένα μήνυμα HAVE⁴, ώστε οι τελευταίοι να αρχίσουν να ζητάνε από τους έχοντες το νέο στοιχείο περιεχομένου που έγινε διαθέσιμο.

³στο BitTorrent, tracker αποκαλείται ο κόμβος που περιέχει και παρέχει πληροφορίες όπως ποιοί κόμβοι κατεβάζουν κάθε στιγμή περιεχόμενο, ποιοί το έχουν ολόκληρο και ποιοί κομμάτια κλπ.

⁴Στο BitTorrent, όταν ένας κόμβος αποκτήσει ένα ολόκληρο τμήμα του αρχείου, στέλνει ένα μήνυμα HAVE στους υπόλοιπους κόμβους, ώστε οι τελευταίοι να γνωρίζουν ότι μπορούν να του το ζητήσουν

3.5 Θεωρητική Ανάλυση Απόδοσης Συστήματος

Σε αυτήν την ενότητα παρουσιάζεται ένα μοντέλο για τη θεωρητική εκτίμηση της επίδρασης που μπορεί να έχει το CCPS σε ένα σύστημα δημοσίευσης/συνδρομής βασισμένο στο περιεχόμενο, όπου τα ερωτήματα αποστέλλονται προς ένα σύνολο κόμβων που δημοσιεύουν το περιεχόμενο, οι οποίοι και αναλαμβάνουν να τα πραγματοποιήσουν και να στείλουν τα αντίστοιχα αποτελέσματα. Εκτός από τη ρεαλιστική περίπτωση της κρυφής μνήμης πεπερασμένης χωρητικότητας, μελετάται επίσης η περίπτωση της άπειρης χωρητικότητας, καθώς μπορεί να υπάρχουν σενάρια όπου η μνήμη είναι αρκετή για να χωρέσει το περιεχόμενο. Κάτι τέτοιο μπορεί να συμβεί όταν οι κόμβοι είναι πολύ καλά οργανωμένοι και υπάρχει μικρός χώρος αναζήτησης και λίγα διαφορετικά ερωτήματα. Επιπλέον, η προσέγγιση της άπειρης μνήμης μπορεί να δώσει εκτίμηση της έκτασης της βελτίωσης που μπορεί να επιφέρει το προτεινόμενο πρωτόκολλο.

Ένα σημαντικό μέρος της ανάλυσης της απόδοσης του CCPS αποτελεί η επιλογή του μοντέλου επαναληπτικότητας των ερωτημάτων. Μολονότι η βιβλιογραφία προτείνει διάφορα μοντέλα για συστήματα βασισμένα σε κλειδιά (π.χ. [27] [62]), στις περιπτώσεις συστημάτων βασισμένων σε λέξεις-κλειδιά, η πιο εμπεριστατωμένη μελέτη είναι αυτή που παρουσιάζεται στο [41]. Εκεί, παρουσιάζεται μια αναλυτική στατιστική παρατήρηση, από την οποία προκύπτει η πιθανότητα επανάληψης ενός ερωτήματος σε συστήματα αναζήτησης με βάση λέξεις-κλειδιά, λαμβάνοντας υπόψη τόσο γενικά όσο και πιο συγκεκριμένα θέματα.

Με βάση την προαναφερθείσα εργασία, το μοντέλο που παρουσιάζεται σε αυτή την ενότητα έχει στόχο να παράσχει μια θεωρητική εκτίμηση ως προς τη βελτίωση του αριθμού των μηνυμάτων ανακάλυψης στο CCPS σε σχέση με ένα τυπικό επιδημικό πρωτόκολλο. Στη συνέχεια, τα αποτελέσματα που προκύπτουν σε αυτήν την ενότητα επιβεβαιώνονται από τις προσομοιώσεις που παρατίθενται στην Ενότητα 3.7.

Ο Πίνακας 1 περιλαμβάνει την επεξήγηση των συμβόλων που θα χρησιμοποιηθούν για τη θεωρητική ανάλυση.

Έστω N_p ο αριθμός των κόμβων του συστήματος που δημοσιεύουν περιεχόμενο. Επιπλέον, έστω p_r η πιθανότητα επανάληψης των ερωτημάτων. Προφανώς, από ένα σύνολο Q ερωτημάτων, τα μοναδικά θα είναι:

$$Q_o = (1 - p_r) \cdot Q \quad (3.1)$$

ενώ τα επαναλαμβανόμενα:

$$Q_r = p_r \cdot Q \quad (3.2)$$

Κάθε δημοσιεύων κόμβος θεωρείται ότι δεσμεύει κάποιο χώρο για να συνδράμει στην κατανεμημένη κρυφή μνήμη. Έστω ότι αυτό το μέγεθος είναι S_c , ίδιο για όλους τους

Πίνακας 1: Επεξήγηση Συμβόλων της Θεωρητικής Ανάλυσης

Επεξήγηση Συμβόλων	
N_p	Αριθμός κόμβων που δημοσιεύουν περιεχόμενο
p_r	Πιθανότητα επανάληψης ερωτήματος
Q	Συνολικός αριθμός ερωτημάτων
Q_o	Αριθμός μοναδικών ερωτημάτων
p_c	Πιθανότητα επιτυχίας στην κρυφή μνήμη
p_{q_i}	Πιθανότητα για κάποιο ερώτημα i να φτάσει σε κάποιο κόμβο του Chord
S_c	Μέγεθος της κρυφής μνήμης σε αριθμό στοιχείων περιεχομένου
$M_d^{(reg)}$	Αριθμός μηνυμάτων ανακάλυψης χωρίς επαναληπτικότητα των ερωτημάτων
$M_d^{(rep-uc)}$	Αριθμός μηνυμάτων ανακάλυψης με επαναληπτικότητα ερωτημάτων και άπειρο μέγεθος κρυφής μνήμης
$M_d^{(rep-lc)}$	Αριθμός μηνυμάτων ανακάλυψης με επαναληπτικότητα ερωτημάτων και πεπερασμένο μέγεθος κρυφής μνήμης
λ	Βελτίωση στον αριθμό των μηνυμάτων ανακάλυψης

κόμβους. Από τη στιγμή που το CCPS κατανέμει το περιεχόμενο στην κρυφή μνήμη με βάση το Chord, το οποίο χρησιμοποιεί συναρτήσεις συνεπούς κατακερματισμού (consistent hashing) [63], δίνοντας έτσι πιθανοτικές εγγυήσεις ομοιόμορφης κατανομής του περιεχομένου σε όλους τους κόμβους του υπερδικτύου, η κρυφή μνήμη ενός κόμβου μπορεί να χρησιμοποιηθεί για τη μελέτη και αναγωγή σε ολόκληρο το σύστημα, λόγω συμμετρίας.

Δεδομένου ότι υπάρχουν μόνο Q_o ερωτήματα, τα οποία μπορούν να επαναληφθούν κατά τη διάρκεια του χρονικού ορίζοντα που εξετάζεται, κάθε ερώτημα θα φτάνει στο σύστημα με πιθανότητα $P_{q_i} = 1/Q_o$. Από τη στιγμή που τα ερωτήματα κατανέμονται ομοιόμορφα στους κόμβους, η πιθανότητα να φτάσει ένα ερώτημα σε κάποιον κόμβο του συστήματος είναι $p_{q_i} = N_p \cdot P_{q_i} = N_p/Q_o$. Συνεπώς, η πιθανότητα για ένα ερώτημα να βρίσκεται σε κάποια θέση στην κρυφή μνήμη του εκάστοτε κόμβου είναι p_{q_i} και, άρα, η πιθανότητα να μη βρίσκεται σε καμία από τις S_c θέσεις είναι $(1 - p_{q_i})^{S_c}$, που σημαίνει ότι η πιθανότητα να βρεθεί το περιεχόμενο σε οποιαδήποτε θέση στην κρυφή μνήμη είναι:

$$p_c = 1 - (1 - p_{q_i})^{S_c} \quad (3.3)$$

και, αντικαθιστώντας το p_{q_i} , η παραπάνω εξίσωση γίνεται:

$$p_c = 1 - \left(1 - \frac{N_p}{(1 - p_r) \cdot Q}\right)^{S_c} \quad (3.4)$$

Στην τυπική περίπτωση της χρήσης επιδημικών πρωτοκόλλων, ο αριθμός των μηνυμάτων ανακάλυψης περιεχομένου ανέρχεται στο συνολικό αριθμό των δημοσιευόντων κόμβων. Για την ακρίβεια, ο αριθμός αυτός είναι στην πράξη μεγαλύτερος, καθώς, λόγω της πιθανοτικής φύσης αυτών των πρωτοκόλλων, μπορεί κάποια μηνύματα να φτάσουν σε κάποιον κόμβο από διαφορετικές διαδρομές. Εντούτοις, για αυτή τη μελέτη, θα θεω-

ρήσουμε ότι επιτυγχάνεται ατομικότητα⁵, χωρίς επιπλέον επιβάρυνση. Έτσι, ο συνολικός αριθμός μηνυμάτων ανακάλυψης θα είναι:

$$M_d^{(reg)} = Q \cdot N_p \quad (3.5)$$

Στην περίπτωση του CCPS, το ερώτημα εφαρμόζεται πρώτα στην κατανεμημένη κρυφή μνήμη και στη συνέχεια, εφόσον δε βρεθεί εκεί, στέλνεται στους δημοσιεύοντες κόμβους. Λαμβάνοντας υπόψη ότι το Chord χρειάζεται $\log N_p$ μηνύματα στη χειρότερη περίπτωση για να εντοπίσει τον κόμβο που μπορεί να φέρει το περιεχόμενο [18], κάθε ερώτημα που γίνεται, θα επιφέρει φόρτο τουλάχιστο $\log N_p$. Στην περίπτωση της κρυφής μνήμης άπειρης χωρητικότητας, όπου το περιεχόμενο θα βρίσκεται πάντα εκτός της πρώτης φοράς, οπότε θα χρειάζεται η χρήση του επιδημικού πρωτοκόλλου με αποστολή του μηνύματος σε όλους τους δημοσιεύοντες κόμβους. Τελικά, ο συνολικός αριθμός των μηνυμάτων είναι:

$$M_d^{(rep-uc)} = Q_o \cdot N_p + Q \cdot \log N_p \quad (3.6)$$

Αντικαθιστώντας με την (3.1), η παραπάνω εξίσωση γίνεται:

$$M_d^{(rep-uc)} = (1 - p_r) \cdot Q \cdot N_p + Q \cdot \log N_p \quad (3.7)$$

Στην περίπτωση της κρυφής μνήμης με περιορισμένο μέγεθος, τα ερωτήματα που υποβάλλονται στους δημοσιεύοντες κόμβους είναι τόσο τα αρχικά, όσο και όλα εκείνα που δε βρέθηκαν στην κρυφή μνήμη, λόγω αντικατάστασής τους. Συνεπώς, ο συνολικός αριθμός μηνυμάτων ανακάλυψης εδώ, είναι:

$$M_d^{(rep-lc)} = (1 - p_c) \cdot Q \cdot N_p + Q \cdot \log N_p \quad (3.8)$$

Σε αυτό το σημείο αξίζει να σημειωθεί ότι η πιθανότητα να βρεθεί περιεχόμενο στην κρυφή μνήμη δεν μπορεί να ξεπεράσει την πιθανότητα επανάληψης του αντίστοιχου ερωτήματος. Ο λόγος είναι ότι, ακόμα και στην καλύτερη περίπτωση της άπειρης χωρητικότητας, θα υπάρχει μία αποτυχία στην αρχή, οπότε και θα μπει το αντίστοιχο περιεχόμενο στην κρυφή μνήμη και στη συνέχεια θα υπάρχει συνεχώς επιτυχία. Όμως, ο λόγος του αριθμού των φορών που επαναλαμβάνεται ένα ερώτημα, προς το συνολικό αριθμό των φορών που πραγματοποιείται είναι στην πραγματικότητα η πιθανότητα επανάληψής του.

Η επίδραση του CCPS στη συνολική απόδοση του συστήματος ως προς τα μηνύματα ανακάλυψης περιεχομένου δίνεται από το λόγο λ , ο οποίος ορίζεται ως εξής:

$$\lambda = 1 - \frac{M_d^{(rep-lc)}}{M_d^{reg}} = p_c - \frac{\log N_p}{N_p} \quad (3.9)$$

⁵Ατομικότητα ορίζεται ως η ιδιότητα ενός συστήματος όπου όλοι οι κόμβοι έχουν λάβει ένα μήνυμα

Δεδομένου ότι η επίδραση του λόγου $[(\log N_p)/N_p]$ μειώνεται όσο μεγαλώνει ο αριθμός N_p των δημοσιευόντων κόμβων $\left(\lim_{N_p \rightarrow \infty} \frac{\log N_p}{N_p} = 0\right)$, η βελτίωση λ ισούται, περίπου, με την πιθανότητα επιτυχίας στην κρυφή μνήμη:

$$\lambda \approx p_c \quad (3.10)$$

Λαμβάνοντας υπόψη ότι η πιθανότητα να βρεθεί περιεχόμενο στην κρυφή μνήμη δεν μπορεί να υπερβεί την πιθανότητα επανάληψης του αντίστοιχου ερωτήματος, η Εξίσωση (3.10) δηλώνει ότι η βελτίωση του CCPS στην απόδοση του συστήματος μπορεί να φτάσει μέχρι την πιθανότητα επανάληψης ερωτημάτων.

3.6 Υλοποίηση Πρωτοκόλλου CCPS

3.6.1 Γενική Προσέγγιση

Η τρέχουσα υλοποίηση του CCPS βασίζεται σε τρία ευρέως διαδεδομένα πρωτόκολλα: το Chord, που χρησιμοποιείται για την κατανεμημένη κρυφή μνήμη και απαρτίζεται από τους δημοσιεύοντες κόμβους, ένα τυπικό επιδημικό πρωτόκολλο, που χρησιμοποιείται για την αποστολή των συνδρομών στους δημοσιεύοντες κόμβους, όταν υπάρχει αποτυχία στην κρυφή μνήμη και σε μια τυπική υλοποίηση του BitTorrent, που χρησιμοποιείται για το διαμοιρασμό των στοιχείων περιεχομένου μεταξύ των συνδρομητών. Σε αυτήν την ενότητα παρουσιάζονται οι βασικές διαδικασίες που λαμβάνουν χώρα για τη δημοσίευση και τη συνδρομή σε πληροφορία. Οι υλοποιήσεις που έγιναν για το CCPS βασίστηκαν στις οδηγίες και τις περιγραφές που δίνονται: για το Chord στο [18], για το επιδημικό πρωτόκολλο στα [50] [64] [65] και για το BitTorrent στα [66] [26].

3.6.2 Δημοσίευση Περιεχομένου

Όπως έχει ήδη εξηγηθεί στην Ενότητα 3.4, το CCPS είναι ένα πρωτόκολλο δημοσίευσης/συνδρομής, όπου οι δημοσιεύοντες κόμβοι εφαρμόζουν τα ερωτήματα που καταφθάνουν τόσο σε υπάρχουσα όσο και σε μελλοντική πληροφορία. Συνεπώς, όταν κάποιος κόμβος δημοσιεύει πληροφορία, θα πρέπει πρώτα να αποθηκεύσει το αντίστοιχο περιεχόμενο στον ενταμιευτή δημοσιεύσεων που διαθέτει, ώστε εκεί να πραγματοποιούνται τα επερχόμενα ερωτήματα, και στη συνέχεια να εφαρμόζει όλα τα ερωτήματα που βρίσκονται αποθηκευμένα στον ενταμιευτή συνδρομών πάνω στο νεοδημοσιευθέν περιεχόμενο. Για κάθε ταίριασμα, ειδοποιείται ο αντίστοιχος συνδρομητής.

Ο Αλγόριθμος 2 περιγράφει τη διαδικασία δημοσίευσης ενός στοιχείου περιεχομένου. Αρχικά, το νέο στοιχείο τοποθετείται στον ενταμιευτή δημοσιεύσεων, ώστε να μπορεί να χρησιμοποιηθεί για το ταίριασμα μελλοντικών συνδρομών. Στη συνέχεια, όλες οι συν-

δρομές που βρίσκονται στον ενταμιευτή συνδρομών εξετάζονται αν απαντώνται από τη νέα δημοσίευση, οπότε ενημερώνεται και ο αντίστοιχος συνδρομητής, όπως επίσης και το σχετικό αρχείο μετα-πληροφορίας, το οποίο αποστέλλεται στον tracker.

Αλγόριθμος 2: Δημοσίευση Περιεχομένου από Κόμβο - PUBLISH_CONTENT

input: content είναι το περιεχόμενο που πρόκειται να δημοσιευτεί

```

1 pub_buffer.put(content);
2 matches[] ← sub_buffer.match(content);
3 forall match ∈ matches[] do
4   | match.subscriber.RECV_CONTENT(match.query, content);
5 end
6 update_metaFile(query);
```

Όταν ένας συνδρομητής λάβει μια νέα δημοσίευση που αντιστοιχεί σε ένα ερώτημά του, την τοποθετεί σε ένα πίνακα κατακερματισμού που αντιστοιχίζει ερωτήματα με τις δημοσιεύσεις που τα απαντούν. Στην πραγματικότητα, αυτός ο πίνακας είναι το περιεχόμενο που διαμοιράζεται μεταξύ των συνδρομητών, με το νέο στοιχείο περιεχομένου να αποτελεί ένα επιπλέον κομμάτι του torrent που αντιπροσωπεύει τη γνώση, η οποία αναπαρίσταται από το ερώτημα της συνδρομής. Ο εν λόγω συνδρομητής είναι και ο πρώτος seeder αυτής της γνώσης, καθώς όλοι οι επόμενοι που εξυπηρετούνται από την κρυφή μνήμη δε θα πάρουν απευθείας απάντηση από τους δημοσιεύοντες κόμβους, αλλά από τους άλλους συνδρομητές του ίδιου ερωτήματος μέσω του πρωτοκόλλου BitTorrent.

Ο Αλγόριθμος 3 περιγράφει τη διαδικασία που λαμβάνει χώρα σε ένα συνδρομητή όταν λάβει ένα στοιχείο περιεχομένου που απαντά σε κάποιο ερώτημα που έχει θέσει στο σύστημα. Το περιεχόμενο αρχικά τοποθετείται στον κατάλληλο ενταμιευτή, ο οποίος σημειώνεται ως ενημερωμένος (αυτή η πληροφορία είναι γενικά χρήσιμη σε εφαρμογές ανώτερου επιπέδου, ώστε να γνωρίζουν ότι έφθασε νέα πληροφορία). Στη συνέχεια ενημερώνει και το αντίστοιχο torrent αρχείο, αν υπάρχει, ζητά την τελευταία λίστα από συνδρομητές της γνώσης που αναπαρίσταται από το εν λόγω ερώτημα και στέλνει σε όλα τα μέλη της το μήνυμα HAVE.

Αλγόριθμος 3: Λήψη Περιεχομένου από Συνδρομητή - RECV_CONTENT

input: query είναι το ερώτημα που πραγματοποίησε ο συνδρομητής
content είναι το στοιχείο περιεχομένου που απαντά το ερώτημα

```

1 pub_table.put(query, content);
2 pub_table.set_updated(query);
3 if have_metaFile(query) then
4   | pieces ← pub_table.get(query).length;
5   | subscribers[] ← update_swarm(query);
6   | forall subscriber ∈ subscribers[] do
7     | subscriber.HAVE(query, pieces);
8   | end
9 end
```

Δεδομένου ότι τα πρωτόκολλα δημοσίευσης/συνδρομής είναι εγγενώς ασύγχρονα, μπορεί ανά πάσα στιγμή να καταφθάσουν νέες δημοσιεύσεις για υπάρχουσες συνδρομές. Ο Αλγόριθμος 3 παρουσιάζει πώς οι συνδρομητές χειρίζονται ένα τέτοιο γεγονός. Εντούτοις, μπορεί να υπάρχουν συνδρομητές, οι οποίοι ήταν εκτός σύνδεσης όταν έγινε η νέα δημοσίευση, αλλά σε κάθε περίπτωση θα πρέπει να ενημερωθούν για το νέο περιεχόμενο. Για το λόγο αυτό, όταν επανέλθουν εντός σύνδεσης, επικοινωνούν με τον tracker ώστε να πάρουν την τελευταία έκδοση του αρχείου μετα-πληροφορίας για τη γνώση που επιθυμούν να λάβουν. Αν κάποιος συνδρομητής διαπιστώσει ότι υπάρχει νέο περιεχόμενο, επικοινωνεί με τη λίστα που παρέλαβε από τον tracker ζητώντας τα κομμάτια που του λείπουν. Η περιοδικότητα αυτής της ενέργειας εξαρτάται από το πόσο συχνά αλλάζει την κατάστασή του ο εκάστοτε συνδρομητής καθώς και το βαθμό κρισιμότητας που έχει θέσει για κάθε ερώτημα. Η διαδικασία ενημέρωσης της γνώσης περιγράφεται στον Αλγόριθμο 4. Συγκεκριμένα, κάθε συνδρομητής ζητά από τον tracker τα torrent αρχεία στα οποία αντιστοιχούν οι συνδρομές του. Για κάθε ένα από αυτά τα αρχεία όπου διαπιστώνει ότι υπάρχουν περισσότερα στοιχεία περιεχομένου από αυτά που έχει ήδη στην κατοχή του ο συνδρομητής, πραγματοποιεί τις αντίστοιχες αιτήσεις στους άλλους συνδρομητές, όπως ακριβώς συμβαίνει και όταν εξυπηρετείται από την κρυφή μνήμη για πρώτη φορά. Κάθε δημοσίευση που καταφθάνει, τοποθετείται στο σχετικό ενταμιευτή.

Αλγόριθμος 4: Ενημέρωση Γνώσης - UPDATE_KNOWLEDGE

```
1 forall query ∈ pub_table.queries do
2   metaFile ← update_tracker(query);
3   for i=query.matches → metaFile.pieces do
4     newMatch ← metaFile.swarm.request(i);
5     pub_table.get(query).put(newMatch);
6   end
7 end
```

3.6.3 Συνδρομή σε Περιεχόμενο

Όπως αναλύεται και στο [16], τα μηνύματα που ανταλλάσσονται στο πλαίσιο ενός σχήματος δημοσίευσης/συνδρομής μπορεί να μεταφέρουν διάφορες πληροφορίες, όπως ένα μοναδικό αναγνωριστικό του περιεχομένου που αιτείται και το συνδρομητή που έκανε την αίτηση. Το προτεινόμενο πρωτόκολλο θεωρεί ότι, εκτός των άλλων πληροφοριών που μπορεί να χρειάζονται, ένα μήνυμα συνδρομής θα πρέπει να μεταφέρει ένα ερώτημα εκπεφρασμένο με λέξεις-κλειδιά, το οποίο ονομάζεται *ερώτημα συνδρομής*.

Όταν ένας συνδρομητής επιθυμεί να θέσει ένα νέο ερώτημα μέσω μιας συνδρομής, αρχικά θα επιχειρήσει να βρει πληροφορία από την κρυφή μνήμη. Αν όντως βρεθεί εκεί, τότε ο συνδρομητής λαμβάνει το αρχείο μετα-πληροφορίας (torrent) που αντιστοιχεί στη γνώση, την οποία αντιπροσωπεύει το ερώτημα, και ξεκινά το διαμοιρασμό της με

τους υπόλοιπους συνδρομητές που την έχουν ήδη λάβει, ή τη λαμβάνουν σε πραγματικό χρόνο, μέσω του πρωτοκόλλου BitTorrent. Αν δε βρεθεί κάτι στην κρυφή μνήμη, τότε είτε το ερώτημα πραγματοποιείται για πρώτη φορά ή έχει αντικατασταθεί στην κρυφή μνήμη. Σε κάθε περίπτωση, ο συνδρομητής θα στείλει το ερώτημα στους δημοσιεύοντες κόμβους χρησιμοποιώντας το επιδημικό πρωτόκολλο. Ο Αλγόριθμος 5 περιγράφει αυτή τη διαδικασία.

Αλγόριθμος 5: Συνδρομή σε Περιεχόμενο - SUBSCRIBE

```
input: subscription είναι η συνδρομή που φέρει το ερώτημα του συνδρομητή
1 caching_peer ← find_successor(subscription.query);
2 metaFile ← caching_peer.get(subscription.query);
3 if metaFile neq null then
4 | share_content(metaFile);
5 else
6 | forall publisher ∈ fanout do
7 | | publisher.RECV_SUBSCRIPTION(subscription);
8 | end
9 end
```

Αρχικά, ο αλγόριθμος χρησιμοποιεί τη διαδικασία `find_successor` του Chord για να βρει ποιος δημοσιεύων κόμβος είναι υπεύθυνος για τη διατήρηση κρυφής μνήμης για το τεθέν ερώτημα. Στη συνέχεια, ο συνδρομητής ζητά από αυτόν τον κόμβο να του στείλει το αρχείο μετα-πληροφορίας, εφόσον το έχει στην κρυφή του μνήμη. Αν το αρχείο υπάρχει, ο τρέχων συνδρομητής θα ξεκινήσει να μοιράζεται τη γνώση με τους υπόλοιπους ενδιαφερόμενους συνδρομητές. Η διαδικασία του διαμοιρασμού ακολουθεί τις αρχές του BitTorrent και, σε γενικές γραμμές, περιλαμβάνει την επικοινωνία με τον tracker ώστε να λάβει τη λίστα με τους ενεργούς συνδρομητές (swarm) και τον αριθμό των στοιχείων περιεχομένου που συναπαρτίζουν τη γνώση, καθώς και τον τρόπο επικοινωνίας με τους υπόλοιπους συνδρομητές. Αν δε βρεθεί το αρχείο στην κρυφή μνήμη, ο συνδρομητής χρησιμοποιεί το επιδημικό πρωτόκολλο ξεκινώντας από ένα σύνολο δημοσιευόντων κόμβων, τους οποίους τοποθετεί στο fanout, για να στείλει τη συνδρομή του.

Όταν ένας δημοσιεύων κόμβος παραλάβει μια συνδρομή, τότε εξάγει από αυτήν το ερώτημα και το πραγματοποιεί πάνω στο περιεχόμενο που έχει δημοσιεύσει. Κάθε στοιχείο περιεχομένου που απαντά στο ερώτημα στέλνεται πίσω στο συνδρομητή μέσω της διαδικασίας `RECV_CONTENT` που περιγράφηκε στον Αλγόριθμο 3. Επιπλέον, η συνδρομή αυτή αποθηκεύεται στον ενταμιευτή συνδρομών, ώστε να είναι διαθέσιμη για ταίριασμα αν στο μέλλον αυτός ο κόμβος πραγματοποιήσει μια νέα δημοσίευση. Τέλος, με βάση και το TTL (time-to-live) του μηνύματος, ο δημοσιεύων κόμβος αποφασίζει αν θα συνεχίσει τη διάχυση του μηνύματος μέσω του επιδημικού αλγορίθμου. Ο Αλγόριθμος 6 περιγράφει την εν λόγω διαδικασία.

Αλγόριθμος 6: Δημοσιεύων Κόμβος που Λαμβάνει Συνδρομή - RECV_SUBSCRIPTION

```

input: subscription είναι η συνδρομή που λαμβάνει ο δημοσιεύων κόμβος
1 sub_table.put(subscription.query);
2 matches[] ← pub_table.match(subscription.query);
3 forall match ∈ matches[] do
4   | subscription.source.RECV_CONTENT(subscription.query, match);
5 end
6 if subscription.TTL > 0 then
7   | subscription.TTL - -;
8   | forall publisher ∈ fanout do
9     | publisher.RECV_SUBSCRIPTION(subscription);
10  | end
11 end

```

3.7 Πειραματική Εκτίμηση Επιδόσεων CCPS

3.7.1 Διαμόρφωση Περιβάλλοντος Προσομοίωσης

Η πειραματική εκτίμηση της απόδοσης του CCPS στοχεύει αφενός στην επιβεβαίωση των αποτελεσμάτων που παρουσιάστηκαν νωρίτερα στην Ενότητα 3.5 και αφετέρου στη μελέτη της συμπεριφοράς του συστήματος ως προς την εξισορρόπηση φόρτου που επιτυγχάνει μεταξύ των συνδρομητών και των κόμβων που δημοσιεύουν περιεχόμενο. Δεδομένου ότι το CCPS σχεδιάστηκε για να βελτιώσει την απόδοση των επιδημικών αλγορίθμων στα πρωτόκολλα δημοσίευσης/εγγραφής βασιζόμενα στο περιεχόμενο, η απόδοσή του θα συγκριθεί με ένα τυπικό επιδημικό πρωτόκολλο, υλοποιημένο με βάση τις σχεδιαστικές αρχές που αναφέρθηκαν στην Ενότητα 3.6.

Το προτεινόμενο πρωτόκολλο έχει υλοποιηθεί στον προσομοιωτή PeerSim [21] [22], ο οποίος έχει χρησιμοποιηθεί εκτενώς για την προσομοίωση πρωτοκόλλων ομότιμων δικτύων. Αποτελείται από μια υλοποίηση ενός τυπικού επιδημικού πρωτοκόλλου, με βάση τις προδιαγραφές που δίνονται στο [65], του πρωτοκόλλου Chord με βάση τις διαδικασίες που ορίζονται στο σχετικό άρθρο [18] και ενός τυπικού πρωτοκόλλου BitTorrent, σύμφωνα με τις οδηγίες που δίνονται στα [66] [26]. Το CCPS αναλαμβάνει την ενορχήστρωση όλων αυτών των επιμέρους πρωτοκόλλων, σύμφωνα με τα όσα παρουσιάστηκαν νωρίτερα στην Ενότητα 3.6.

Σε ό,τι αφορά το μοντέλο επανάληψης των ερωτημάτων, χρησιμοποιήθηκαν διάφορες προσεγγίσεις. Αρχικά, ακολουθώντας και τις παρατηρήσεις που έγιναν στο [41], θεωρήθηκε ένας συνολικός πληθυσμός από ερωτήματα. Στη συνέχεια, με βάση την πιθανότητα επανάληψης ενός ερωτήματος, δημιουργήθηκε το σύνολο των μοναδικών ερωτημάτων. Έτσι, κάθε ερώτημα που υποβάλλεται στο σύστημα εξάγεται ομοιόμορφα από τα μοναδικά.

Ωστόσο, πέραν του προαναφερθέντος μοντέλου, έγιναν επιπλέον προσομοιώσεις

θεωρώντας ότι η πιθανότητα επανάληψης των ερωτημάτων ακολουθεί το νόμο του Zipf. Στόχος αυτής της προσέγγισης ήταν να εξεταστεί η απόδοση της κρυφής μνήμης με αυτήν την κατανομή, οπότε διερευνήθηκε μόνο η πιθανότητα επιτυχίας στην κρυφή μνήμη καθώς και ο αριθμός των μηνυμάτων ανακάλυψης όταν εφαρμόζονται τα δύο διαφορετικά μοντέλα. Ακολουθώντας την ίδια λογική με προηγουμένως, και σε αυτήν την περίπτωση δημιουργήθηκε ένα σύνολο με διαφορετικά ερωτήματα, από το οποίο επιλεγόταν κάθε φορά ένα ερώτημα, με βάση τον επόμενο αριθμό Zipf. Τα πειράματα εκτελέστηκαν 100 φορές έκαστο και στις δύο περιπτώσεις.

Τα σενάρια που θεωρήθηκαν περιελάμβαναν διαφορετικό αριθμό από δημοσιεύοντες κόμβους και ερωτήματα, ώστε να εξεταστεί η ικανότητα κλιμάκωσης του πρωτοκόλλου κάτω από διάφορες συνθήκες. Συγκεκριμένα, στις επόμενες ενότητες θα παρουσιαστούν πειράματα που περιλαμβάνουν 4800, 9600, 14400 και 19200 δημοσιεύοντες κόμβους και 15000, 30000, 45000 και 6000 ερωτήματα που επαναλαμβάνονται με πιθανότητα από 0 μέχρι 0,8 με βήμα 0,2. Σε όλες αυτές τις περιπτώσεις θεωρήθηκε άπειρο μέγεθος κρυφής μνήμης, προκειμένου να εξεταστεί η επίδοση του πρωτοκόλλου σε ιδανικές συνθήκες.

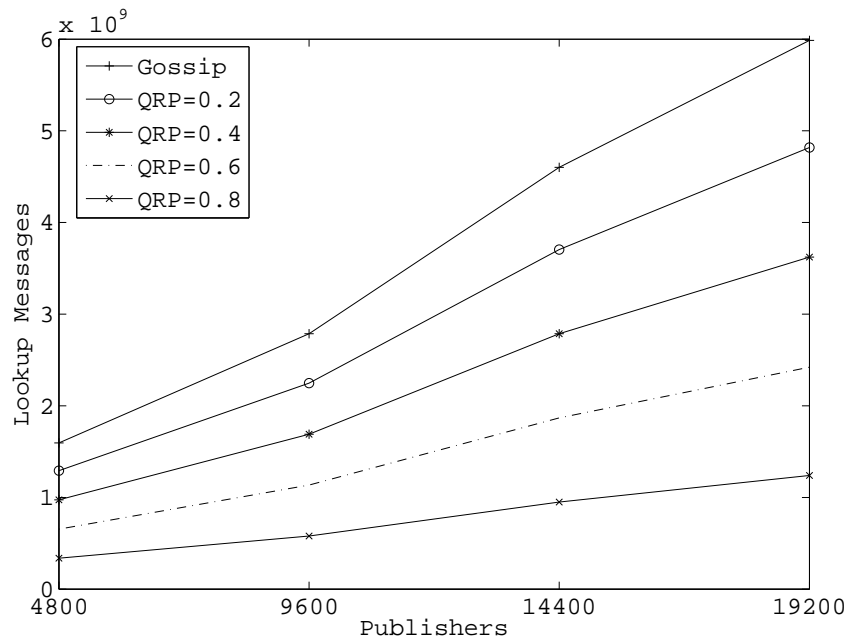
Στη συνέχεια, πραγματοποιήθηκαν προσομοιώσεις με περιορισμένο μέγεθος κρυφής μνήμης, ώστε να επιβεβαιωθεί πειραματικά η Εξίσωση 3.10. Για το σκοπό αυτό, θεωρήθηκε σύστημα με 1000 δημοσιεύοντες κόμβους, 1500 συνδρομητές, 200000 ερωτήματα, πιθανότητα επανάληψης ερωτημάτων 0.8 και μέγεθος κρυφής μνήμης από 5 μέχρι 50 αρχεία μετα-πληροφορίας, με βήμα 5. Στην περίπτωση που χρησιμοποιήθηκε ο νόμος του Zipf, η εκθετική παράμετρος τέθηκε ενδεικτικά σε 0,82.

Εκτός του κόστους επικοινωνίας που υφίστανται οι δημοσιεύοντες κόμβοι από την αποστολή των ερωτημάτων, οι προσομοιώσεις διερευνούν επίσης κατά πόσο επιτυγχάνεται εξισορρόπηση φόρτου στο σύστημα, μέσω της ενεργής συμμετοχής των συνδρομητών στην αποστολή του περιεχομένου. Επιπλέον, μετρήθηκε η αποτελεσματικότητα της χρήσης εύρους ζώνης όταν δε χρησιμοποιείται επιδημικό πρωτόκολλο, το οποίο, λόγω της πιθανοτικής του φύσης, μπορεί να προκαλέσει αποστολή του ίδιου μηνύματος περισσότερες της μία φορές σε κάποιον κόμβο. Μάλιστα, αυτό το φαινόμενο έχει αντίκτυπο και στον αριθμό των μηνυμάτων που στέλνει ο δημοσιεύων κόμβος, καθώς μπορεί να στείλει το ίδιο στοιχείο περιεχομένου παραπάνω από μία φορές, κάθε φορά που λαμβάνει ένα αντίστοιχο ερώτημα.

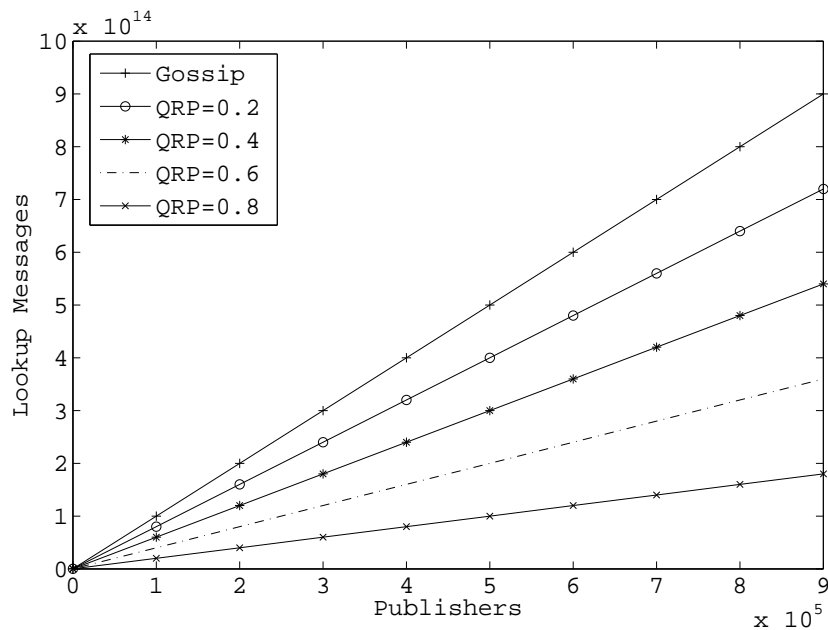
Σημειώνεται ότι στις επόμενες ενότητες, όπου αναφέρεται πιθανότητα επανάληψης ερωτημάτων ίση με το μηδέν, υπονοείται η χρήση επιδημικού πρωτοκόλλου (βλ. Σχήματα 35, 36, 38 και 39).

3.7.2 Μηνύματα Ανακάλυψης Περιεχομένου

Στην περίπτωση της κρυφής μνήμης άπειρης χωρητικότητας, είναι προφανές από τις προσομοιώσεις ότι το CCPS εκμεταλλεύεται πλήρως (όπως άλλωστε αποδείχθηκε και



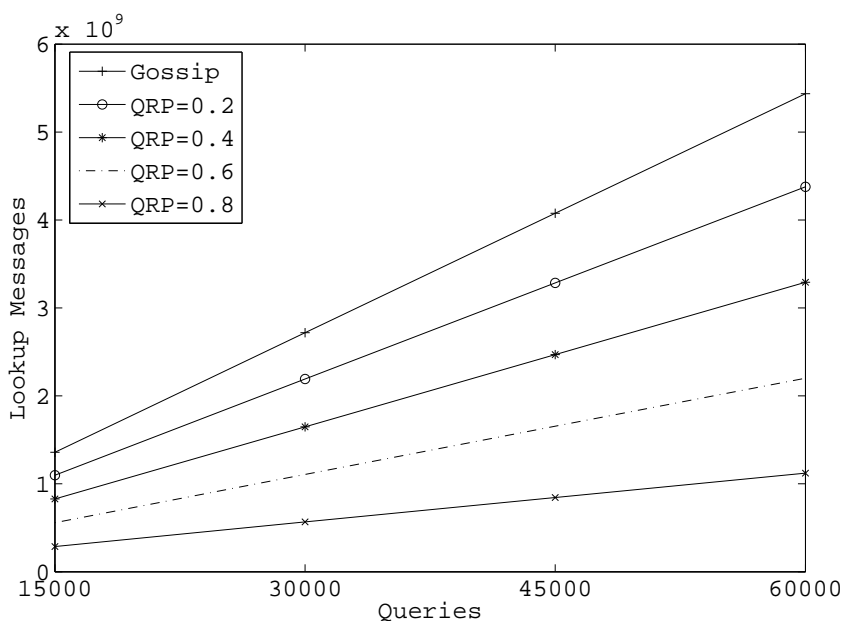
Σχήμα 26: Μηνύματα Ανακάλυψης ως προς Δημοσιεύοντες Κόμβους (Προσομοιώσεις)



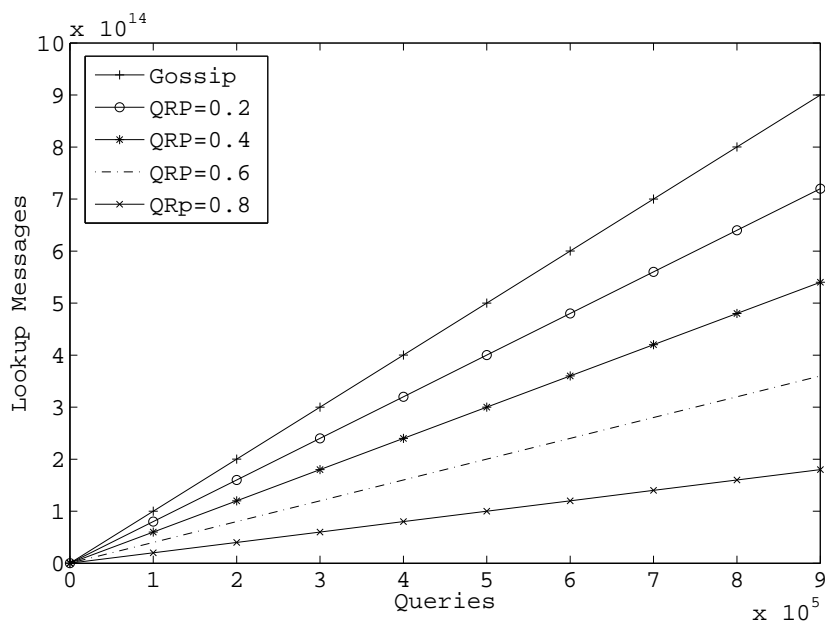
Σχήμα 27: Μηνύματα Ανακάλυψης ως προς Δημοσιεύοντες Κόμβους (Ανάλυση)

στη θεωρητική ανάλυση του πρωτοκόλλου) την επαναληπτικότητα των ερωτημάτων. Οι προσομοιώσεις που παρατίθενται μετρούν τον αριθμό των μηνυμάτων ανακάλυψης περιεχομένου, όσο το σύστημα μεγαλώνει τόσο σε αριθμό κόμβων όσο και ερωτημάτων.

Το Σχήμα 26 παρουσιάζει τον αριθμό μηνυμάτων ανακάλυψης στο σύστημα για 4800, 9600, 14400 και 19200 κόμβους, με την πιθανότητα επανάληψης ερωτημάτων να αυξά-



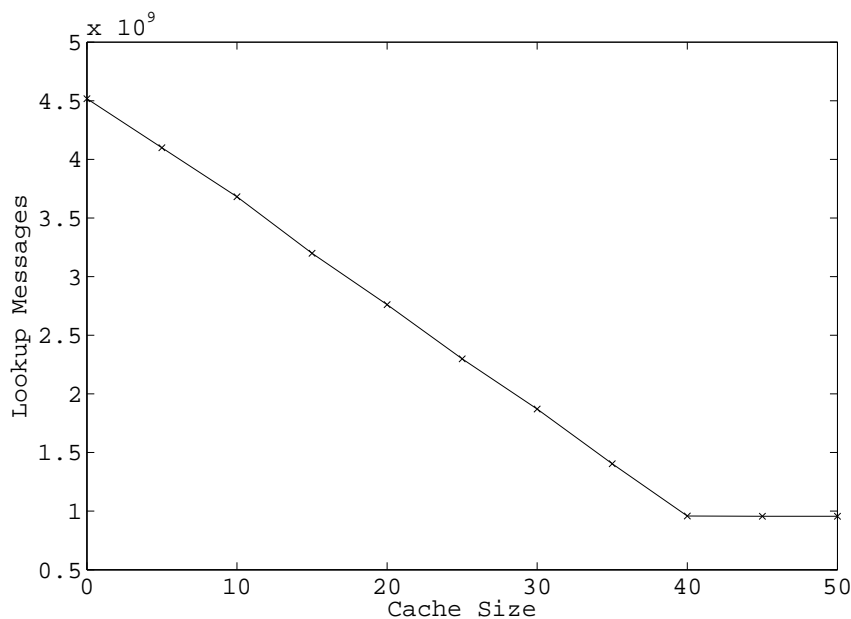
Σχήμα 28: Μηνύματα Ανακάλυψης ως προς Ερωτήματα (Προσομοιώσεις)



Σχήμα 29: Μηνύματα Ανακάλυψης ως προς Ερωτήματα (Ανάλυση)

νεται με βήμα 0,2 από 0 σε 0,8. Δεδομένου ότι τα επιδημικά πρωτόκολλα αντιστοιχούν στη μηδενική πιθανότητα επανάληψης, φαίνεται ότι η επίδραση του CCPS στο σύστημα γίνεται ολοένα σημαντικότερη όσο αυξάνεται ο αριθμός των κόμβων. Το Σχήμα 27 απεικονίζει τα αντίστοιχα αριθμητικά αποτελέσματα από τη θεωρητική ανάλυση, επιβεβαιώνοντας ότι η τελευταία συμφωνεί με τα αποτελέσματα των προσομοιώσεων.

Ωστόσο, η επίδραση του πρωτοκόλλου στην επιβάρυνση του συστήματος δε φαίνε-

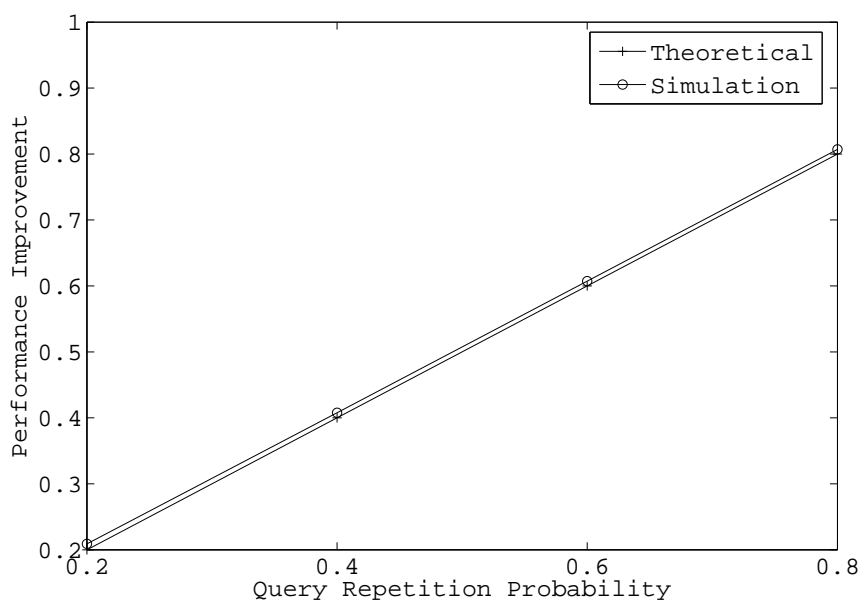


Σχήμα 30: Μηνύματα Ανακάλυψης για 200000 Ερωτήματα και Περιορισμένη Κρυφή Μνήμη

ται μόνο ως προς τον αριθμό των κόμβων, αλλά και των ερωτημάτων που τίθενται. Από τη στιγμή που το CCPS εκμεταλλεύεται την επαναληπτικότητα των ερωτημάτων, είναι αναμενόμενο να παρουσιάζει βελτίωση ως προς το συνολικό αριθμό τους. Αυτό φαίνεται ξεκάθαρα στο Σχήμα 26, το οποίο δείχνει τη μείωση του αριθμού των μηνυμάτων ανακάλυψης με την αύξηση της πιθανότητας επανάληψης ερωτημάτων. Φυσικά, ο αριθμός των μηνυμάτων εξακολουθεί να αυξάνεται, ωστόσο ο ρυθμός αύξησης είναι μικρότερος όσο μεγαλύτερη είναι και η πιθανότητα επανάληψης. Η έκταση του φαινομένου απεικονίζεται και στο Σχήμα 29 που περιλαμβάνει τα αριθμητικά αποτελέσματα με βάση τη θεωρητική ανάλυση.

Στην περίπτωση της περιορισμένης κρυφής μνήμης, θεωρήθηκε σύστημα με 1000 δημοσιεύοντες κόμβους και πιθανότητα επανάληψης ερωτημάτων ίση με 0,8. Το Σχήμα 30 παρουσιάζει την εξέλιξη του αριθμού των μηνυμάτων ανακάλυψης όταν η χωρητικότητα της κρυφής μνήμης αυξάνεται από 0 (που αντιστοιχεί στην περίπτωση ενός τυπικού επιδημικού πρωτοκόλλου) σε 50 αρχεία μετα-πληροφορίας. Εδώ πρέπει να σημειωθεί ότι σύμφωνα με το θεωρητικό μοντέλο του συστήματος, όταν το μέγεθος της κρυφής μνήμης ανά κόμβο ξεπεράσει το 40, τότε παύει να υπάρχει βελτίωση, καθώς δεν υπάρχει άλλο περιεχόμενο να αποθηκευτεί. Αυτό αποδεικνύεται και από τις προσομοιώσεις, όπως φαίνεται στο Σχήμα 30.

Η βελτίωση λ στην επίδοση του συστήματος ως προς την επιβάρυνση των μηνυμάτων ανακάλυψης, την οποία επιφέρει η χρήση του CCPS, έχει εκτιμηθεί θεωρητικά από την Εξίσωση (3.10). Στην περίπτωση της κρυφής μνήμης άπειρης χωρητικότητας, το λ αναμέ-

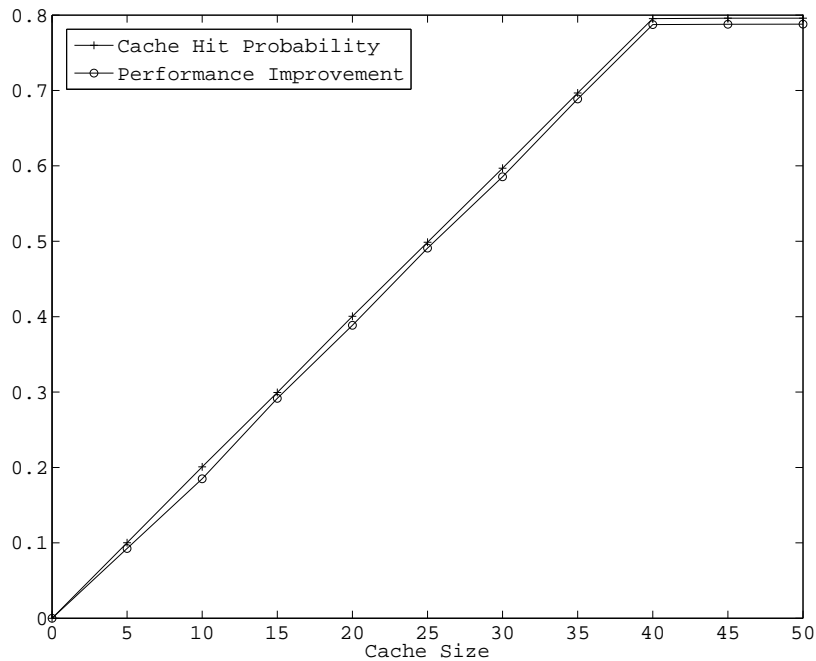


Σχήμα 31: Βελτίωση λ στην Επίδοση του Συστήματος (Απειρη Κρυφή Μνήμη)

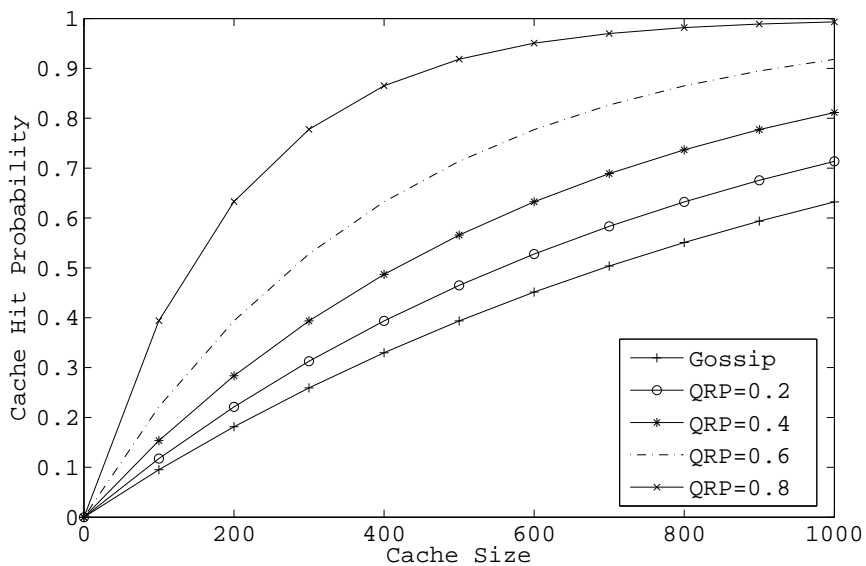
νεται να είναι ίσο με την πιθανότητα επανάληψης των ερωτημάτων. Ωστόσο, εξαιτίας της επικάλυψης των μηνυμάτων (για την οποία γίνεται ειδικά αναφορά στην Ενότητα 3.7.4), η κατάσταση στην πράξη είναι ακόμα καλύτερη. Το Σχήμα 31 υποδεικνύει τη γραμμική αύξηση της βελτίωσης με την πιθανότητα επανάληψης, όπως υποστηρίχθηκε και από τη θεωρητική ανάλυση.

Θεωρώντας περιορισμένη κρυφή μνήμη, τα αποτελέσματα είναι αντίστοιχα. Έτσι, όπως φαίνεται και στο Σχήμα 32, όσο αυξάνεται το μέγεθος της κρυφής μνήμης, αυξάνεται και το λ . Ξανά, όταν το μέγεθος ξεπεράσει το 40, οπότε δεν υπάρχει αντικατάσταση περιεχομένου, η επίδοση παραμένει σταθερή. Επιπλέον, η βελτίωση φαίνεται ότι είναι περίπου ίση με την πιθανότητα εύρεσης περιεχομένου στην κρυφή μνήμη, όπως προκύπτει και από την Εξίσωση (3.10). Ειδικά ως προς την πιθανότητα εύρεσης περιεχομένου στην κρυφή μνήμη, το Σχήμα 33 παρουσιάζει την εξέλιξή της με βάση αριθμητικές τιμές και την ανάλυση. Για πιθανότητα επανάληψης ερωτημάτων ίση με 0,8, η οποία θεωρήθηκε και στις προσομοιώσεις, είναι φανερό ότι η συμπεριφορά μεταξύ της προσομοίωσης και της ανάλυσης είναι αντίστοιχη.

Ολοκληρώνοντας τη μελέτη του συστήματος ως προς τα μηνύματα ανακάλυψης, πραγματοποιήθηκαν προσομοιώσεις όπου η πιθανότητα επανάληψης των ερωτημάτων ακολουθεί το νόμο του Zipf, τα αποτελέσματα των οποίων παρατίθενται στο Σχήμα 34. Η εν λόγω γραφική παράσταση, αφενός επιβεβαιώνει τον ισχυρισμό της Εξίσωσης (3.10) και, αφετέρου, αποδεικνύει ότι το μοντέλο είναι ανεξάρτητο της κατανομής που ακολουθεί η πιθανότητα επανάληψης των ερωτημάτων. Μια επιπλέον παρατήρηση που προκύπτει από αυτό το σχήμα είναι ότι η επίδοση της κρυφής μνήμης αυξάνεται σε αυτήν την περίπτωση,

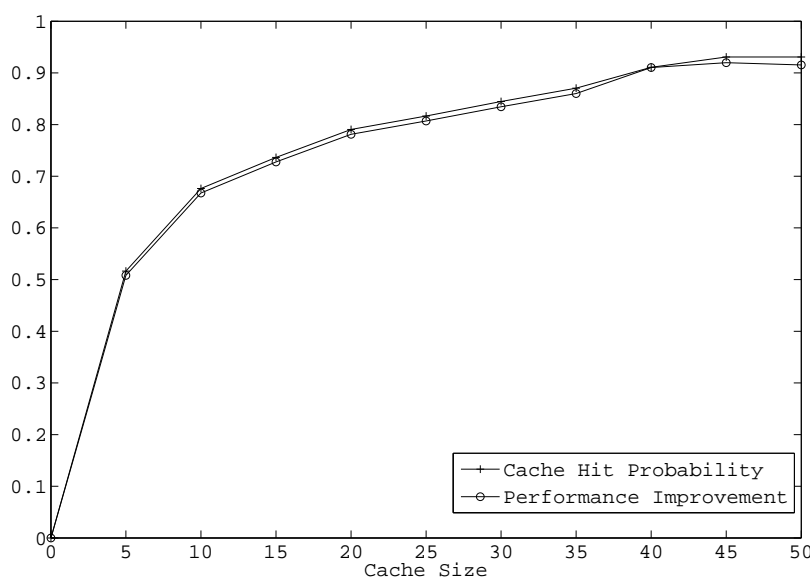


Σχήμα 32: Βελτίωση λ στην Επίδοση του Συστήματος με Περιορισμένη Κρυφή Μνήμη (Προσομοιώσεις)



Σχήμα 33: Πιθανότητα Εύρεσης Περιεχομένου στην Κρυφή Μνήμη (Θεωρητική Ανάλυση)

κάτι που εξηγείται λόγω της μεγαλύτερης δημοφιλίας κάποιων ερωτημάτων, πράγμα που οδηγεί σε μεγαλύτερη χρήση της κρυφής μνήμης.

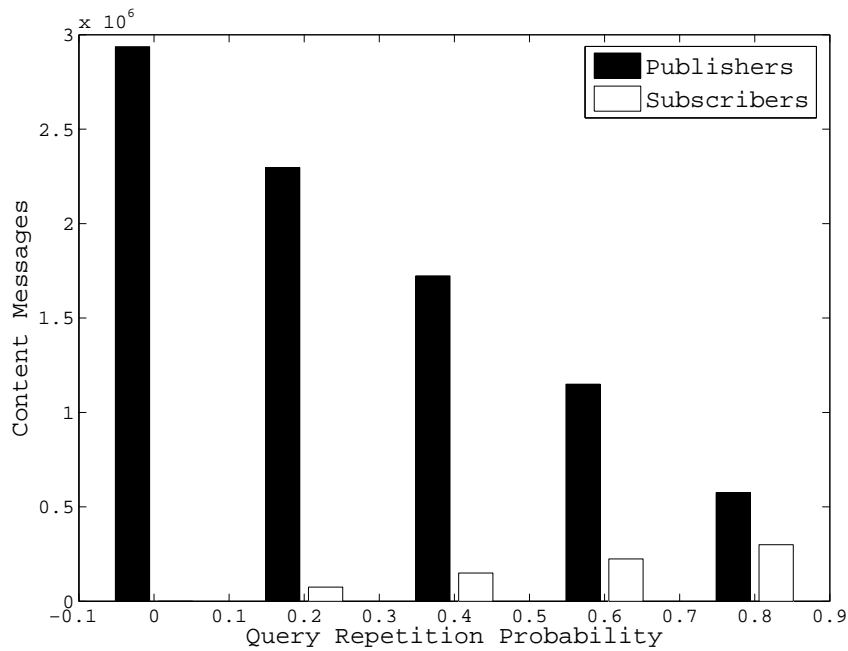


Σχήμα 34: Βελτίωση λ στην Απόδοση του Συστήματος για Zipf Κατανομή

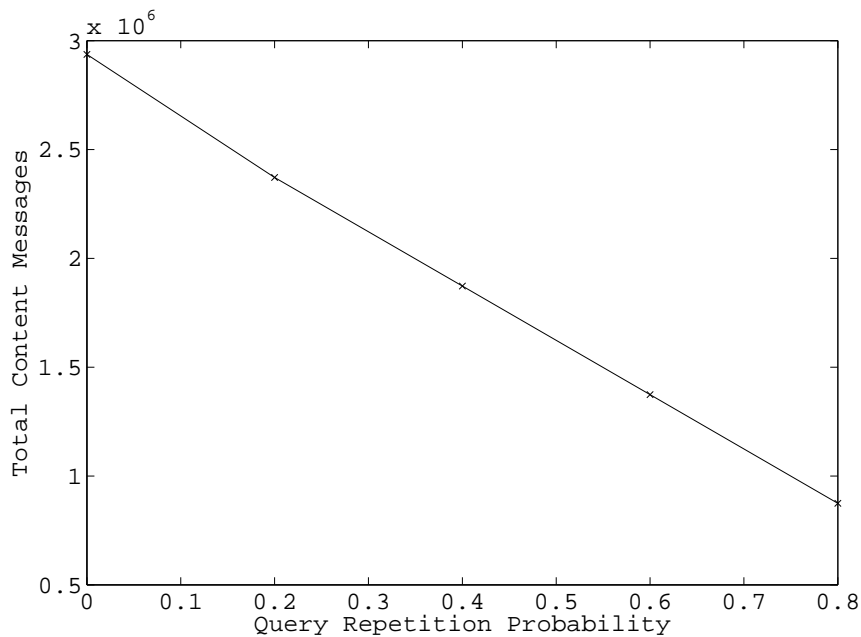
3.7.3 Μηνύματα Περιεχομένου

Ένα χαρακτηριστικό του CCPS είναι ότι επιδρά θετικά και στην εξισορρόπηση φόρτου του συστήματος με το να αποδεσμεύει του δημοσιεύοντες κόμβους από το να στέλνουν εκείνοι όλες τις απαντήσεις στους συνδρομητές. Έτσι, μέσω της αναπαράστασης του περιεχομένου που αντιστοιχίζεται σε ένα ερώτημα ως στοιχεία γνώσης, την οποία διαμοιράζονται οι αντίστοιχοι συνδρομητές χρησιμοποιώντας το πρωτόκολλο BitTorrent, το CCPS επιτυγχάνει μεγαλύτερη κατανομή του φόρτου επικοινωνίας στους κόμβους του συστήματος. Τα αποτελέσματα των προσομοιώσεων που παρατίθενται σε αυτήν την ενότητα επιβεβαιώνουν αυτόν τον ισχυρισμό.

Το Σχήμα 35 παρουσιάζει τον αριθμό των μηνυμάτων περιεχομένου που στέλνουν οι δημοσιεύοντες κόμβοι και οι συνδρομητές για διαφορετικές πιθανότητες επανάληψης ερωτημάτων. Όπως φαίνεται και στο σχήμα, ο αριθμός των μηνυμάτων περιεχομένου που στέλνουν οι δημοσιεύοντες κόμβοι μειώνονται όσο αυξάνεται η πιθανότητα επανάληψης ερωτημάτων, ενώ αντίθετα τα μηνύματα που στέλνουν οι συνδρομητές αυξάνονται. Μάλιστα, όταν η πιθανότητα επανάληψης ερωτημάτων είναι ίση με 0,8, τα μηνύματα περιεχομένου που στέλνουν οι συνδρομητές φτάνουν περίπου το 50% των μηνυμάτων που στέλνουν οι δημοσιεύοντες κόμβοι. Διαβάζοντας την ίδια γραφική παράσταση από μια διαφορετική γωνία, μια ακόμα χρήσιμη παρατήρηση είναι ότι ο ρυθμός αύξησης των μηνυμάτων περιεχομένου που στέλνουν οι συνδρομητές αυξάνεται κατά 75%, όταν ο αντίστοιχος ρυθμός μείωσης των μηνυμάτων περιεχομένου που στέλνουν οι δημοσιεύοντες κόμβοι μειώνεται κατά 50%. Το συμπέρασμα από αυτήν την παρατήρηση είναι ότι με την αύξηση της πιθανότητας επανάληψης ερωτημάτων, αυξάνεται και η συνεργατική συμπεριφορά των



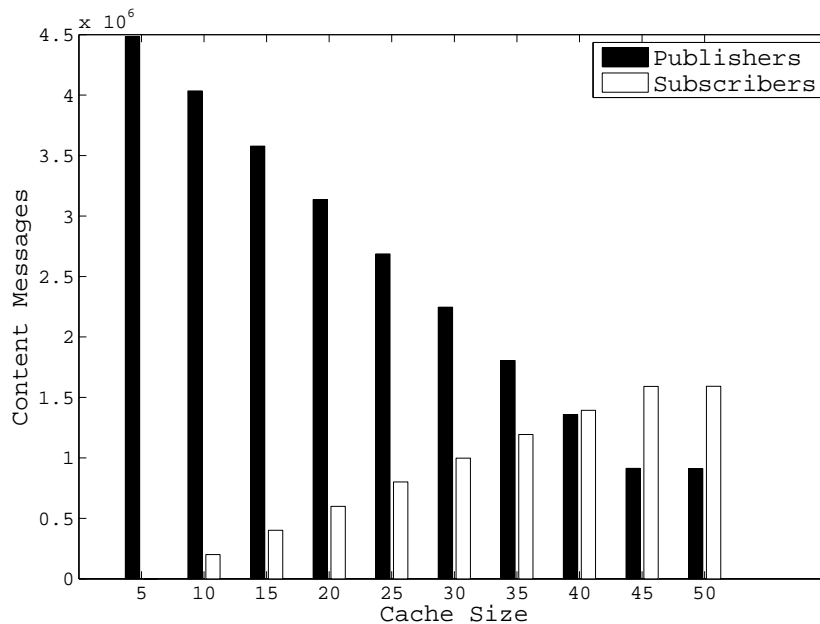
Σχήμα 35: Εξισορρόπηση Φόρτου στα Μηνύματα Περιεχομένου



Σχήμα 36: Εξέλιξη Μηνυμάτων Περιεχομένου (Απειρη Κρυφή Μνήμη)

κόμβων, πράγμα που οδηγεί και σε καλύτερη συνολική επίδοση. Αυτό φαίνεται ξεκάθαρα στο Σχήμα 36, το οποίο απεικονίζει το συνολικό αριθμό μηνυμάτων περιεχομένου στο σύστημα για διαφορετικές πιθανότητες επανάληψης ερωτημάτων.

Οι παραπάνω παρατηρήσεις επιβεβαιώνονται και στην περίπτωση της περιορισμέ-

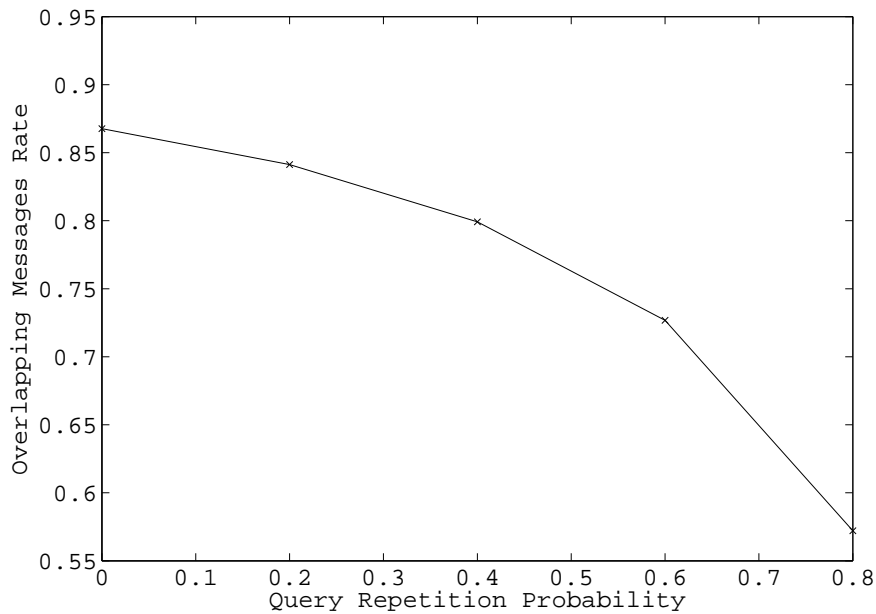


Σχήμα 37: Μηνύματα Περιεχομένου για 200.000 Ερωτήματα και Περιορισμένη Κρυφή Μνήμη

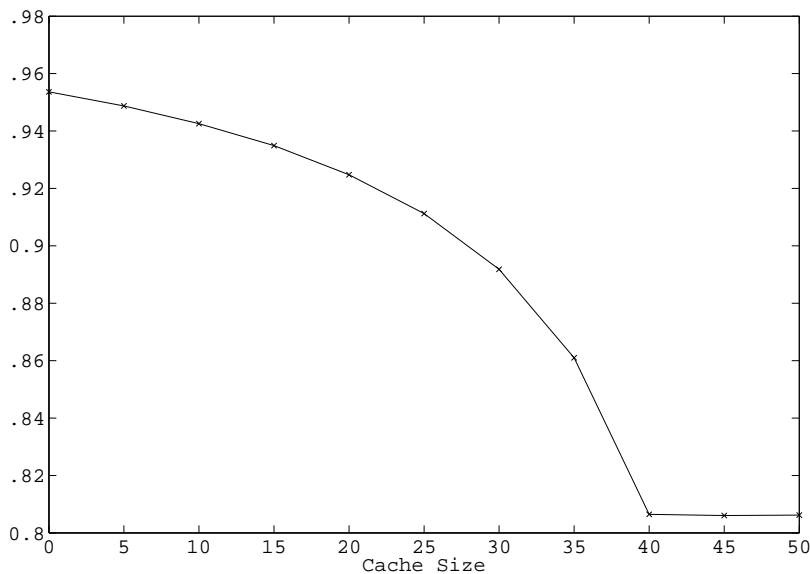
νης χωρητικότητας της κρυφής μνήμης. Επιπλέον, λαμβάνοντας υπόψη ότι η παρακολούθηση του συστήματος σε αυτήν την περίπτωση αφορά περισσότερα ερωτήματα, το Σχήμα 37 δείχνει πως όταν το μέγεθος της κρυφής μνήμης γίνει 40, ο αριθμός των μηνυμάτων που στέλνουν οι συνδρομητές γίνεται μεγαλύτερος από αυτόν των δημοσιευόντων κόμβων. Το γεγονός αυτό αποδεικνύει ότι όσο μεγαλώνει ο αριθμός των ερωτημάτων, τόσο βελτιώνεται η επίδοση του συστήματος, πάντα, βέβαια, σε συνάρτηση και με το μέγεθος της κρυφής μνήμης. Όπως ήταν αναμενόμενο, όταν το μέγεθος της κρυφής μνήμης ξεπεράσει το 40, δεν υπάρχει περαιτέρω βελτίωση στο σύστημα, δεδομένου ότι σε αυτό το μέγεθος έχει εξαντληθεί η ικανότητά της, μιας και φιλοξενεί όλα τα ερωτήματα.

3.7.4 Επικάλυψη Μηνυμάτων

Ένα μειονέκτημα των επιδημικών πρωτοκόλλων, οφειλόμενο στην πιθανοτική τους φύση, είναι ότι ένα μήνυμα μπορεί να φτάσει σε ένα κόμβο πολλές φορές επειδή ακολουθεί διαφορετικά μονοπάτια. Ένας τρόπος για να περιοριστεί αυτό το φαινόμενο είναι να προστεθεί κάποιο πεδίο στην επικεφαλίδα του μηνύματος, το οποίο να περιλαμβάνει το μονοπάτι που έχει ακολουθήσει μέχρι εκείνη τη στιγμή, έτσι ώστε να αποκλειστούν οι εν λόγω κόμβοι ως πιθανοί επόμενοι προορισμοί. Ακόμα και έτσι, όμως, δεν είναι δυνατόν να εξαλειφθεί το φαινόμενο, καθώς τελικά ο κάθε κόμβος αποφασίζει ανεξάρτητα από τους υπόλοιπους. Κατά συνέπεια, μπορεί κόμβοι να λάβουν επικαλυπτόμενες τέτοιες αποφάσεις.



Σχήμα 38: Επικαλυπτόμενα Μηνύματα με Άπειρη Κρυφή Μνήμη



Σχήμα 39: Επικαλυπτόμενα Μηνύματα με Περιορισμένη Κρυφή Μνήμη

Το Σχήμα 38 απεικονίζει τον αριθμό των μηνυμάτων ανακάλυψης σε σχέση με την πιθανότητα επανάληψης ερωτημάτων. Όπως φαίνεται, στην περίπτωση του τυπικού επιδημικού αλγορίθμου (πιθανότητα μηδέν) η επικάλυψη είναι μεγαλύτερη του 85%. Ωστόσο, όσο αυξάνεται η πιθανότητα επανάληψης ερωτημάτων, οπότε μειώνεται η χρήση του επιδημικού αλγορίθμου, η επίδραση του φαινομένου μειώνεται. Στην περίπτωση της κρυφής μνήμης περιορισμένου μεγέθους και για πιθανότητα επανάληψης ερωτημάτων, η συμπε-

ριφορά του συστήματος φαίνεται στο Σχήμα 39. Ξανά, όταν το μέγεθος της κρυφής μνήμης γίνει 40, η επίδραση παραμένει σταθερή, αφού το CCPS έχει πιάσει τη μέγιστη απόδοση.

Κεφάλαιο 4

Υποστήριξη Ετερογένειας με τα Πρότυπα MPEG-21 και MPEG-M

4.1 Βασικές Δομές Δεδομένων του MPEG-21

Το MPEG-21[67] αποτελεί ένα σύνολο από πρότυπα, τα οποία επικεντρώνονται στον καθορισμό ενός ενιαίου, πρότυπου τρόπου με τον οποίο μπορεί να περιγραφεί η πληροφορία. Για το σκοπό αυτό, ορίζει τη θεμελιώδη δομή του ψηφιακού αντικειμένου (Digital Item - DI) [68], ενός πλαισίου, το οποίο μπορεί να περιλάβει πληροφορία που σχετίζεται με κάποιον πόρο.

Φυσικά, εκτός από το ψηφιακό αντικείμενο αυτό καθεαυτό, το MPEG-21 ορίζει και τρόπους για να περιγραφεί ένας πόρος που περιέχεται σε ένα ψηφιακό αντικείμενο, για να προστατευτούν τα πνευματικά δικαιώματα των δημιουργών, για να παρασχεθούν διαφορετικοί τρόποι πρόσβασης στον πόρο, αναπαράστασής του στη συσκευή του χρήστη καθώς και ένα σύνολο από άλλες επιπλέον δυνατότητες, οι οποίες περιγράφονται στα 20 τεχνικά κείμενα που αποτελούν το MPEG-21.

4.1.1 Το Ψηφιακό Αντικείμενο

Το πρότυπο του ψηφιακού αντικειμένου καθορίζει ένα XML μοντέλο με βάση το οποίο μπορεί να αναπαρασταθεί οποιοσδήποτε πόρος. Για το σκοπό αυτό παρέχεται ένα πλήθος από XML στοιχεία, τα οποία μπορούν να ενθυλακώσουν όλες τις απαραίτητες πληροφορίες που απαιτούνται. Όπως αναφέρεται και στο πρότυπο, σκοπός του δεν είναι να ορίσει μια συγκεκριμένη γλώσσα αναπαράστασης πόρων, παρά να παράσχει ένα ενιαίο σύνολο από αφηρημένους όρους και έννοιες που μπορούν να χρησιμοποιηθούν για αυτήν την αναπαράσταση [68]. Στην παρούσα ενότητα δε θα αναφερθούν παρά μόνο εκείνα τα σημεία του προτύπου, τα οποία χρησιμοποιούνται από το καταναεμημένο σύστημα αρ-

χειών EuterpeFS. Τα σχήματα XML που παρατίθενται προέρχονται από σχετικό σύνδεσμο της ISO [69].

Το υψηλότερο επίπεδο οργάνωσης ενός ψηφιακού αντικειμένου αποτελεί το στοιχείο Container. Ένα στοιχείο Container μπορεί να περιέχει κανένα ή περισσότερα επιπλέον στοιχεία Container, κανένα ή περισσότερα στοιχεία Item και κανένα ή περισσότερα στοιχεία Descriptor. Σημασιολογικά, το Container σκοπό έχει μόνο να υποδείξει ότι ξεκινά μια ενότητα του ψηφιακού αντικειμένου, χωρίς να δίνει κάποιο άλλο στοιχείο για το περιεχόμενο αυτό καθαυτό. Το σχήμα XML για το στοιχείο Container είναι το ακόλουθο:

```

1 <element name="Container" type="didl:ContainerType"
2     substitutionGroup="didmodel:Container"/>
3 <complexType name="ContainerType">
4     <complexContent>
5         <extension base="didmodel:ContainerType">
6             <sequence>
7                 <element ref="didmodel:Descriptor" minOccurs="0"
8                     maxOccurs="unbounded"/>
9                 <element ref="didmodel:Container" minOccurs="0"
10                    maxOccurs="unbounded"/>
11                <element ref="didmodel:Item" minOccurs="0"
12                    maxOccurs="unbounded"/>
13            </sequence>
14            <attributeGroup ref="didl:ID_ATTRS"/>
15            <anyAttribute namespace="##other"
16                processContents="lax"/>
17        </extension>
18    </complexContent>
19 </complexType>

```

Το θεμελιώδες στοιχείο για την ενθυλάκωση των πληροφοριών που αφορούν ένα πόρο (π.χ. αρχείο) από ένα ψηφιακό αντικείμενο είναι το Item. Ένα Item μπορεί να συμπεριλαμβάνει ένα σύνολο από άλλα στοιχεία Item ή στοιχεία Component. Επιπλέον, μπορεί να περιλαμβάνονται και στοιχεία Descriptor, Choice, Condition και Annotation, όπως περιγράφεται και στο αντίστοιχο σχήμα XML:

```

1 <element name="Item" type="didl:ItemType"
2     substitutionGroup="didmodel:Item"/>
3 <complexType name="ItemType">
4     <complexContent>
5         <extension base="didmodel:ItemType">
6             <sequence>
7                 <element ref="didmodel:Condition" minOccurs="0"
8                     maxOccurs="unbounded"/>
9                 <element ref="didmodel:Descriptor" minOccurs="0"
10                    maxOccurs="unbounded"/>

```

```

11     <element ref="didmodel:Choice" minOccurs="0"
12           maxOccurs="unbounded"/>
13     <choice minOccurs="0" maxOccurs="unbounded">
14       <element ref="didmodel:Item"/>
15       <element ref="didmodel:Component"/>
16     </choice>
17     <element ref="didmodel:Annotation" minOccurs="0"
18           maxOccurs="unbounded"/>
19   </sequence>
20   <attributeGroup ref="didl:ID_ATTRS"/>
21   <anyAttribute namespace="##other"
22         processContents="lax"/>
23 </extension>
24 </complexContent>
25 </complexType>

```

Το στοιχείο `Descriptor` παρέχει πληροφορίες, οι οποίες περιγράφουν το στοιχείο στο οποίο περιέχεται. Αυτές οι πληροφορίες μπορούν να είναι είτε δομημένες είτε αδόμητες, με το πρότυπο να παρέχει την ευελιξία να χρησιμοποιηθούν διάφορα σχήματα μεταδεδομένων. Το σχήμα XML για ένα στοιχείο `Descriptor` είναι το ακόλουθο:

```

1 <element name="Descriptor" type="didl:DescriptorType"
2       substitutionGroup="didmodel:Descriptor"/>
3 <complexType name="DescriptorType">
4   <complexContent>
5     <extension base="didmodel:DescriptorType">
6       <sequence>
7         <element ref="didmodel:Condition" minOccurs="0"
8               maxOccurs="unbounded"/>
9         <element ref="didmodel:Descriptor" minOccurs="0"
10              maxOccurs="unbounded"/>
11       <choice>
12         <element ref="didmodel:Component"/>
13         <element ref="didmodel:Statement"/>
14       </choice>
15     </sequence>
16     <attributeGroup ref="didl:ID_ATTRS"/>
17     <anyAttribute namespace="##other" processContents="lax"/>
18   </extension>
19 </complexContent>
20 </complexType>

```

Το στοιχείο `Component` δίνει έναν πόρο (στοιχείο `Resource`) με την περιγραφή (στοιχείο `Descriptor`). Σύμφωνα με το πρότυπο, η περιγραφή αυτή αφορά μεταδεδομένα που σχετίζονται με δομικά στοιχεία του πόρου (π.χ. πληροφορίες για αλγόριθμο κρυπτογράφησης) και όχι με πληροφορία που αφορά το ίδιο το περιεχόμενο του πόρου. Το σχήμα XML του στοιχείου `Component` δίνεται παρακάτω:

```

1 <element name="Component" type="didl:ComponentType"
2     substitutionGroup="didmodel:Component"/>
3 <complexType name="ComponentType">
4   <complexContent>
5     <extension base="didmodel:ComponentType">
6       <sequence>
7         <element ref="didmodel:Condition" minOccurs="0"
8             maxOccurs="unbounded"/>
9         <element ref="didmodel:Descriptor" minOccurs="0"
10            maxOccurs="unbounded"/>
11        <element ref="didmodel:Resource" maxOccurs="unbounded"/>
12        <element ref="didmodel:Anchor" minOccurs="0"
13            maxOccurs="unbounded"/>
14      </sequence>
15      <attributeGroup ref="didl:ID_ATTRS"/>
16      <anyAttribute namespace="##other" processContents="lax"/>
17    </extension>
18  </complexContent>
19 </complexType>

```

Όπως φαίνεται και από το προαναφερθέν σχήμα, ένα στοιχείο Descriptor μπορεί να περιέχει ένα στοιχείο Statement. Το τελευταίο χρησιμοποιείται για να συμπεριλάβει ένα τμήμα κειμένου, το οποίο όμως αποτελεί ένα καθαρό συμπληρωματικό κομμάτι του πόρου και όχι έναν πόρο αυτό καθαυτό. Έτσι, για παράδειγμα, ένα Statement μπορεί να περιγράφει το λειτουργικό σύστημα για το οποίο είναι προορισμένη μια εφαρμογή που αναπαρίσταται από το ψηφιακό αντικείμενο. Το σχήμα XML για το στοιχείο Statement είναι το ακόλουθο:

```

1 <element name="Statement" type="didl:StatementType"
2     substitutionGroup="didmodel:Statement"/>
3 <complexType name="StatementType" mixed="true">
4   <complexContent mixed="true">
5     <extension base="didmodel:StatementType">
6       <sequence>
7         <any namespace="##any" processContents="lax" minOccurs="0"/>
8       </sequence>
9       <attribute name="mimeType" type="string" use="required"/>
10      <attribute name="ref" type="anyURI"/>
11      <attribute name="encoding" type="string"/>
12      <attribute name="contentEncoding" type="NMTOKENS"/>
13      <anyAttribute namespace="##other" processContents="lax"/>
14    </extension>
15  </complexContent>
16 </complexType>

```

Τέλος, το στοιχείο Resource αποτελεί το πλέον ουσιαστικό για το καταναμημένο

σύστημα αρχείων EuterpeFS, καθώς είναι αυτό που περιλαμβάνει το URI του πόρου. Εδώ πρέπει να καταστεί σαφές ότι ένα ψηφιακό αντικείμενο περιλαμβάνει μια αναφορά σε έναν πόρο και όχι τον πόρο αυτό καθαυτό. Το σχήμα XML για το στοιχείο Resource είναι το παρακάτω:

```

1 <element name="Resource" type="didl:ResourceType"
2     substitutionGroup="didmodel:Resource"/>
3 <complexType name="ResourceType" mixed="true">
4     <complexContent mixed="true">
5         <extension base="didmodel:ResourceType">
6             <sequence>
7                 <any namespace="##any" processContents="lax" minOccurs="0"/>
8             </sequence>
9             <attribute name="mimeType" type="string" use="required"/>
10            <attribute name="ref" type="anyURI"/>
11            <attribute name="encoding" type="string"/>
12            <attribute name="contentEncoding" type="NMTOKENS"/>
13            <anyAttribute namespace="##other" processContents="lax"/>
14        </extension>
15    </complexContent>
16 </complexType>

```

4.1.2 Η Γλώσσα Περιγραφής Δικαιωμάτων

Ειδικά για να οριστούν τα δικαιώματα πρόσβασης σε έναν πόρο, το MPEG-21 ορίζει τη γλώσσα περιγραφής δικαιωμάτων (Rights Expression Language - REL) [70]. Η γλώσσα αυτή είναι σε θέση να περιγράψει ένα σύνολο κανόνων που καθορίζουν εφόσον κάποιος χρήστης ή ομάδα χρηστών έχουν το δικαίωμα να εξασκήσουν ένα σύνολο δικαιωμάτων πάνω στον πόρο στον οποίο αναφέρονται. Η REL παρέχει τη δυνατότητα για πολύπλοκες εκφράσεις εξουσιοδότησης, οι οποίες δεν αναλύονται μέσα στο πλαίσιο της παρούσας εργασίας και ο αναγνώστης παραπέμπεται στο ίδιο το πρότυπο [70]. Εντούτοις, θα παρουσιαστεί το βασικό πλαίσιο μέσα στο οποίο έχει στηθεί το σύστημα εξουσιοδότησης της REL.

Η REL μπορεί να χρησιμοποιηθεί τόσο για να εκφράσει τα δικαιώματα που θα πρέπει να υφίστανται για κάποιον συγκεκριμένο πόρο, ώστε να γίνεται έλεγχος πρόσβασης κάθε φορά που επιχειρείται η προσπέλασή του, όσο και να εκφράσει τα δικαιώματα που έχει η οντότητα που επιχειρεί να προσπελάσει τον πόρο. Έτσι, ένα τυπικό σχήμα εξουσιοδότησης βασισμένο στη REL έχει ως εξής: Ο δημιουργός ενός πόρου (π.χ. ενός τραγουδιού που διατίθεται μέσω ενός αρχείου MP3 από μια συγκεκριμένη υπηρεσία) φτιάχνει με τη REL ένα περίγραμμα άδειας χρήσης στην οποία ορίζει ότι μόνο όποιος έχει πληρώσει στην εταιρεία X μπορεί να έχει πρόσβαση στον πόρο. Τότε, ο απλός χρήστης που επιθυμεί να ακούσει το τραγούδι θα πληρώσει την εταιρεία X και θα λάβει από αυτήν μια άδεια χρήσης του εν λόγω τραγουδιού. Τέλος, ο απλός χρήστης θα ζητήσει από την προαναφερθείσα

υπηρεσία το τραγούδι, δίνοντάς της την άδεια χρήσης του. Η υπηρεσία επιβεβαιώνει την εγκυρότητα της άδειας και, αναλόγως, επιτρέπει ή όχι το είδος της πρόσβασης στον πόρο, το οποίο περιγράφει η αρχική άδεια του δημιουργού.

Η δομή της REL είναι αρκετά απλή στη βάση της. Αυτό που ορίζει είναι ένα σύνολο από σχήματα XML με τα οποία μπορεί να εκφραστεί το δικαίωμα (Right) μιας οντότητας (Principal) πάνω σε κάποιον πόρο (Resource). Φυσικά, μέσα σε αυτές τις απλές εκφράσεις, μπορούν να ληφθούν υπόψη και άλλες παράμετροι που έχουν να κάνουν με το πλαίσιο εξουσιοδότησης, να δημιουργηθούν ομάδες εξουσιοδότησης κλπ.

4.1.3 Το Λεξικό Δικαιωμάτων

Η γλώσσα περιγραφής δικαιωμάτων, μπορεί να συνδυαστεί με το λεξικό δικαιωμάτων (Rights Data Dictionary - RDD) [71], ένα επεκτάσιμο λεξικό των πιθανών δικαιωμάτων που μπορεί να εξασκήσει ένας χρήστης πάνω σε κάποιον πόρο. Σε αυτό το λεξικό περιλαμβάνονται οι πρότυποι όροι που αφορούν αυτά τα δικαιώματα, οι σχέσεις μεταξύ τους καθώς και οι επεξηγήσεις τους. Αυτό το λεξικό μπορεί να επεκταθεί και με επιπλέον όρους, οι οποίοι, βέβαια, θα πρέπει να έχουν προσυμφωνηθεί με όλα τα εμπλεκόμενα μέρη, αλλά σε κάθε περίπτωση υπάρχει ένα κοινό σημείο αναφοράς στο RDD και στον τρόπο χρήσης αυτών των όρων στη REL.

4.2 Οι Τεχνολογίες του MPEG-M

Το MPEG-M (MPEG Middleware) [72] αποτελείται από πέντε τεχνικά κείμενα, ενώ η δεύτερη έκδοση του έχει προχωρήσει προς την τελική δημοσίευση των προτύπων κατά τα δύο τελευταία χρόνια. Σκοπός αυτής της τεχνολογίας είναι να ορίσει την αρχιτεκτονική, το περιβάλλον διεπαφής, το λογισμικό, τις υπηρεσίες, τον τρόπο σύνθεσης των υπηρεσιών καθώς και τις κατευθυντήριες για τη χρήση των παραπάνω τεχνολογιών, του μεσισμικού του MPEG. Πρόκειται για ένα μεσισμικό γενικού σκοπού, το οποίο στηρίζεται κατεξοχήν στις δομές δεδομένων που ορίζει MPEG-21.

Η χρήση μιας πρότυπης πλατφόρμας μεσισμικού επιτρέπει την ενοποίηση ενός καταναμημένου συστήματος, όπου οι χρήστες είναι σε θέση να επιλέξουν υπάρχουσες ή ακόμα και να προχωρήσουν σε δικές τους υλοποιήσεις των, κατά τα άλλα, αυστηρά καθορισμένων πρωτοκόλλων. Αυτή η προσέγγιση δίνει τη δυνατότητα στο καταναμημένο σύστημα αρχείων EutereFS, που παρουσιάζεται στο Κεφάλαιο 5, να υποστηρίξει το διαμοιρασμό περιεχομένου από μεγάλο εύρος χρηστών, οι οποίοι μπορεί να χρησιμοποιούν διαφορετικές πλατφόρμες και αρχιτεκτονικές στις συσκευές τους.

Τα συστατικά μέρη του μεσισμικού είναι αυστηρά καθορισμένα και διακριτά μεταξύ τους, ώστε να είναι δυνατή αφενός η επαναχρησιμοποίησή τους σε διαφορετικές πε-

ριπτώσεις και αφετέρου ο συνδυασμός τους με ξεχωριστές υλοποιήσεις κάθε φορά. Έτσι, ακολουθώντας μια λογική προσανατολισμένη σε υπηρεσίες, το πρότυπο καθορίζει έναν αριθμό θεμελιωδών υπηρεσιών, που καλύπτουν ένα μεγάλο φάσμα ενεργειών, από δημιουργία, περιγραφή και αναζήτηση περιεχομένου, μέχρι αυθεντικοποίηση και εξουσιοδότηση χρηστών για πρόσβαση σε αυτό.

Στα πλαίσια της παρούσας διατριβής έγιναν συνεισφορές στο δεύτερο μέρος του, που αφορά τις προγραμματιστικές διεπαφές [73] [74] [75] [76], στο τρίτο μέρος του, που αφορά τον τρόπο υλοποίησης [77] [78] [79] [80] [81] [82] [83], στο τέταρτο μέρος του, που αφορά τα πρωτόκολλα και τις στοιχειώδεις υπηρεσίες που παρέχονται [84] [85] και στο πέμπτο μέρος του που αφορά το συνδυασμό των στοιχειωδών υπηρεσιών του προτύπου [86] για τη δημιουργία ενοποιημένων υπηρεσιών.

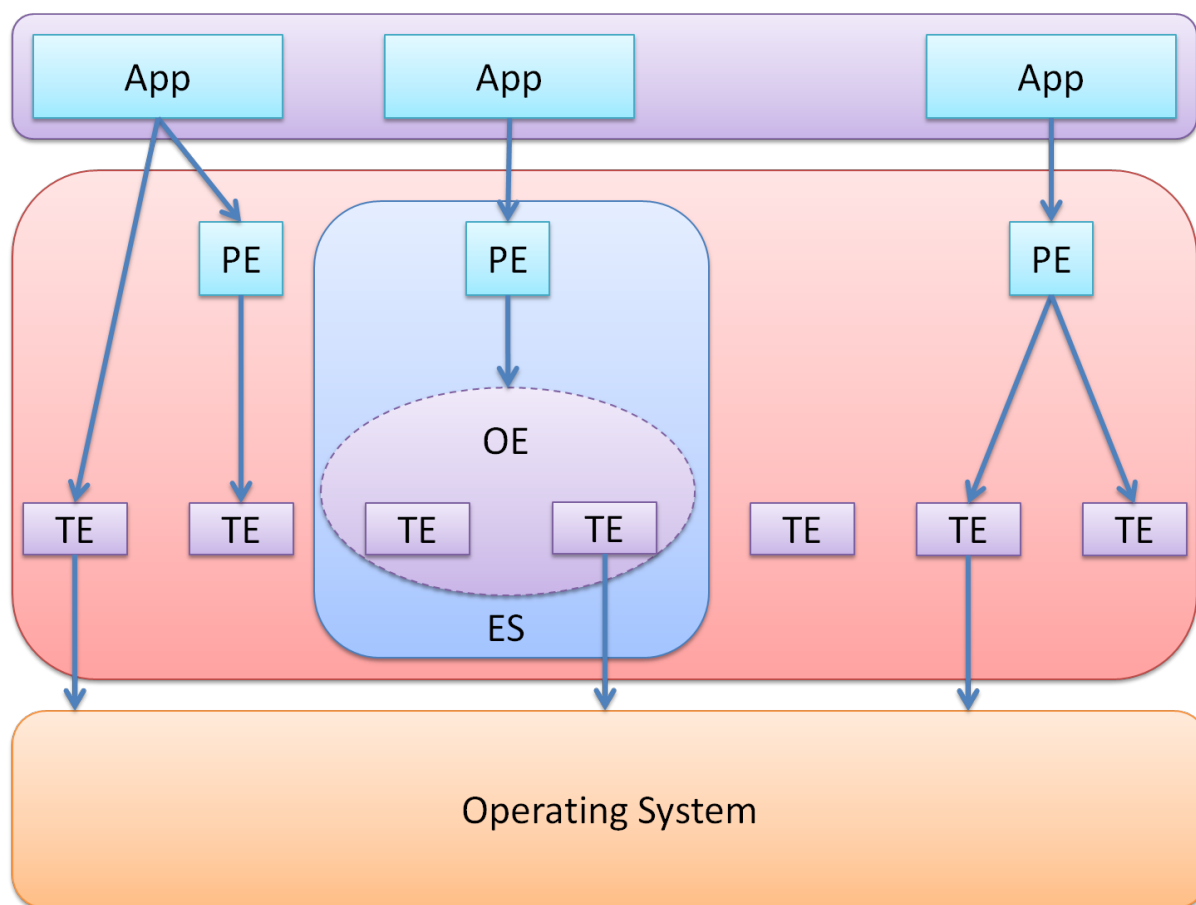
4.2.1 Αρχιτεκτονική του MPEG-M

Η αρχιτεκτονική του MPEG-M [87] επιτρέπει στις εκάστοτε εφαρμογές να χρησιμοποιήσουν τα επιμέρους συστατικά του, τόσο μεμονωμένα όσο και ενορχηστρωμένα με διάφορους τρόπους. Επιπλέον, η λειτουργικότητά του παρέχεται τόσο σε επίπεδο προγραμματιστικών διεπαφών, οι οποίες υλοποιούνται σε επίπεδο βιβλιοθηκών, όσο και ολοκληρωμένων υπηρεσιών, οι οποίες περιλαμβάνουν έτσι κι αλλιώς και τις απαραίτητες βιβλιοθήκες.

Το Σχήμα 40 δίνει μια αφαιρετική εικόνα της αρχιτεκτονικής του MPEG-M. Τα θεμελιώδη συστατικά του στοιχεία είναι οι μηχανές πρωτοκόλλων (Protocol Engines - PEs) και οι μηχανές τεχνολογίας (Technology Engines - TEs). Οι πρώτες χειρίζονται τα πρωτόκολλα που έχουν καθοριστεί στο πρότυπο και τα οποία έχουν να κάνουν με την επικοινωνία με τις αποκαλούμενες στοιχειώδεις υπηρεσίες (Elementary Services - ESs), οι οποίες θα περιγραφούν σε επόμενη ενότητα. Οι μηχανές τεχνολογίας αφορούν συγκεκριμένες, χαμηλού επιπέδου διαδικασίες. Με δεδομένη την αυστηρά καθορισμένη σε πολύ συγκεκριμένα πλαίσια λειτουργικότητα των μηχανών τεχνολογίας, η οποία συνήθως οδηγεί στην ανάγκη για χρήση περισσότερων της μιας για την εκπλήρωση μιας διαδικασίας υψηλότερου επιπέδου, έχει εισαχθεί η έννοια της μηχανής ενορχήστρωσης (Orchstrator Engine - OE), η οποία έχει συγκεκριμένη προγραμματιστική διεπαφή και αναλαμβάνει να συνδυάσει περισσότερες της μιας μηχανές τεχνολογίας για να επιστρέψει το τελικό αποτέλεσμα.

Έτσι, όπως φαίνεται και στο Σχήμα 40, μια εφαρμογή μπορεί να χρησιμοποιήσει είτε απευθείας μια μηχανή τεχνολογίας, είτε εμμέσως μέσω ενός πρωτοκόλλου που υλοποιεί μια μηχανή πρωτοκόλλου. Επιπλέον, μπορεί μέσω της τελευταίας να προσπελάσει μια στοιχειώδη υπηρεσία, η οποία ενθυλακώνει την αντίστοιχη μηχανή πρωτοκόλλου, καθώς και τις απαραίτητες μηχανές τεχνολογίας, οι οποίες συνδυάζονται μέσω μιας μηχανής ενορχήστρωσης.

Σε αυτό το σημείο αξίζει να σημειωθεί ότι κάποιες μηχανές τεχνολογίας μπορεί να χρειάζονται βιβλιοθήκες του λειτουργικού συστήματος, προκειμένου να εκπληρώσουν τη



Σχήμα 40: Αρχιτεκτονική του Μεσισμικού MPEG-M

λειτουργικότητά τους. Αν και αυτό μπορεί να περιορίσει το πεδίο εφαρμογής του, πολλές φορές δεν μπορεί να αποφευχθεί ιδιαίτερα όταν αφορά επικοινωνία με συσκευές. Ένα τέτοιο χαρακτηριστικό παράδειγμα αποτελεί η μηχανή τεχνολογίας ασφάλειας, μια υλοποίηση της οποίας μπορεί να κάνει χρήση αναγνώστη έξυπνων καρτών για κρυπτογράφηση και ψηφιακή υπογραφή.

4.2.2 Μηχανές Πρωτοκόλλου και Τεχνολογίας του MPEG-M

Οι μηχανές πρωτοκόλλου και τεχνολογίας του MPEG-M ορίζονται στο δεύτερο μέρος του προτύπου [88] μαζί με τις προγραμματιστικές διεπαφές τους (Application Programming Interfaces - APIs). Ο ρόλος τους είναι να παρέχουν πρόσβαση στις διαθέσιμες τεχνολογίες με ενιαίο τρόπο, έτσι ώστε να μπορεί να υποστηριχθούν διαφορετικές υλοποιήσεις, με διαφανή τρόπο ως προς τις εφαρμογές υψηλότερου επιπέδου.

Το βασικό τους κομμάτι, το οποίο είναι μάλιστα οριζόντιο σε όλες τις μηχανές, αφορά τη διαχείριση των σχημάτων που τους αντιστοιχούν, τόσο σε επίπεδο πρωτοκόλλων, όσο και σε επίπεδο δομών δεδομένων. Έτσι, παρέχουν προγραμματιστικές διεπαφές

για δημιουργία, πρόσβαση και επεξεργασία της πληροφορίας που τις αφορά. Ειδικά οι μηχανές τεχνολογίας, λόγω και του περιορισμένου πεδίου εφαρμογής της καθεμιάς, μπορούν να συνδυαστούν από μηχανές ενορχήστρωσης, με το πρότυπο να περιλαμβάνει τις μηχανές ενορχήστρωσης που απαιτήθηκαν για τις στοιχειώδεις υπηρεσίες.

Οι μηχανές πρωτοκόλλου έχουν στόχο την υλοποίηση της επικοινωνίας μεταξύ των εφαρμογών υψηλότερου επιπέδου και των στοιχειωδών υπηρεσιών. Συνεπώς, υπάρχει σχέση ένα-προς-ένα μεταξύ τους, ήτοι κάθε μηχανή πρωτοκόλλου αντιστοιχίζεται σε μια στοιχειώδη υπηρεσία. Ρόλος τους είναι να λαμβάνουν και να αποκωδικοποιούν το μήνυμα της εφαρμογής πελάτη, να καλούν τις αντίστοιχες μηχανές τεχνολογίας (ή ενορχήστρωσης) και να επιστρέφουν την κατάλληλη απάντηση, με βάση το πρότυπο σχήμα επικοινωνίας της εκάστοτε στοιχειώδους υπηρεσίας.

Η προγραμματιστική διεπαφή που παρέχουν οι μηχανές πρωτοκόλλου δύναται να χρησιμοποιηθεί τόσο στην πλευρά του εξυπηρετητή της υπηρεσίας όσο και του πελάτη, δίνοντας τη δυνατότητα πρόσβασης στις αντίστοιχες δομές μηνυμάτων με έναν ενιαίο τρόπο και, μέσω της πρότυπης υλοποίησης που δίνεται στο τρίτο μέρος του προτύπου [89], απαλλάσσοντας τελικά τον πελάτη από την επιβάρυνση της επεξεργασίας των μηνυμάτων. Ειδικά για την κλήση της διαδικασίας που πραγματοποιείται στον εξυπηρετητή, η προγραμματιστική διεπαφή των μηχανών πρωτοκόλλου παρέχουν μια μέθοδο, η οποία έχει ξεχωριστές υλοποιήσεις, ανάλογα με το περιβάλλον, όπου εκτελείται: για τον εξυπηρετητή, καλούνται οι κατάλληλες μηχανές τεχνολογίας, ενώ για τον πελάτη, καλείται η πλευρά του εξυπηρετητή να πράξει τις κατάλληλες ενέργειες. Με αυτόν τον τρόπο, δίνεται η δυνατότητα δημιουργίας εφαρμογών, ανεξάρτητα από το αν το περιβάλλον όπου θα εκτελούνται θα περιλαμβάνει τις τεχνολογικές βιβλιοθήκες που χρειάζονται ή αν θα τις καλεί από μια απομακρυσμένη τοποθεσία.

Οι μηχανές τεχνολογίας αφορούν τη διαχείριση των τεχνολογιών που είναι απαραίτητες προκειμένου να εκπληρώσουν το ρόλο τους οι στοιχειώδεις υπηρεσίες. Η πλειοψηφία τους δε, αφορά στο χειρισμό των XML σχημάτων δομών δεδομένων, κυρίως του MPEG-21, αλλά και άλλων προτύπων όπου χρειάζεται (π.χ. XML ψηφιακές υπογραφές, XML κρυπτογράφηση και MPEG-7), πάνω στις οποίες βασίζεται ή εφαρμόζεται η εκάστοτε τεχνολογία. Ξεκινώντας αυτά τα δεδομένα, οι προγραμματιστικές διεπαφές των μηχανών τεχνολογίας αποτελούνται στην πλειοψηφία τους από χειριστές των σχημάτων που χρησιμοποιούν, ενώ μπορεί να συνοδεύονται και από διεπαφές που αφορούν συγκεκριμένες λειτουργίες που πρέπει να εκπληρώσουν.

Όπως είναι προφανές από τα παραπάνω, το πρότυπο έχει χτιστεί έχοντας κυρίως ως οδηγό την πληροφορία που διακινείται, παρέχοντας έναν ενιαίο τρόπο πρόσβασης σε αυτήν. Εντούτοις, για κάποιες μηχανές όπου υπήρξαν απαιτήσεις γενικού σκοπού για πιο συγκεκριμένες τεχνολογίες, περιλήφθηκαν και διεπαφές που να τις ικανοποιούν. Φυσικά, οι υλοποιήσεις τους δεν είναι απαραίτητο να αφορούν το σύνολο των διεπαφών, παρά μόνο αυτές που μπορεί να χρειάζονται κατά περίπτωση.

Σε ό,τι αφορά την ενορχήστρωση των μηχανών τεχνολογίας, το πρότυπο δε δίνει συγκεκριμένο τρόπο για να γίνεται, δεδομένου ότι πρόκειται για λειτουργίες χαμηλού επιπέδου. Εντούτοις, ορίζει μία μηχανή ενορχήστρωσης για κάθε στοιχειώδη υπηρεσία, σε πλήρη αντιστοιχία και με τις διεπαφές των μηχανών πρωτοκόλλου, οι οποίες έχουν διττό ρόλο: αφενός, παρέχουν ένα δοκιμασμένο τρόπο που μπορεί να ακολουθηθεί και για άλλες υλοποιήσεις και αφετέρου, διευκολύνουν τη δημιουργία μιας στοιχειώδους υπηρεσίας, παρέχοντας προγραμματιστικές διεπαφές για όλα τα στάδια της υλοποίησής της (μηχανή πρωτοκόλλου, μηχανές τεχνολογίας και ενορχήστρωση).

Έχοντας δημιουργήσει ένα πρότυπο με τις προγραμματιστικές διεπαφές των μηχανών που χρησιμοποιούνται από τις εφαρμογές, όπως αναφέρθηκε και παραπάνω, δίνεται η δυνατότητα για εναλλακτικές υλοποιήσεις ίδιων μηχανών, οι οποίες μπορεί να χρησιμοποιούνται ακόμα και στα πλαίσια μιας μοναδικής εφαρμογής. Για να υποστηριχθεί αυτό με έναν ομογενή τρόπο, το πρότυπο καθορίζει και το XML σχήμα όπου δηλώνονται ρητά οι υλοποιήσεις καθώς και οι παράμετροι των μηχανών που πρόκειται να χρησιμοποιηθούν. Ένα τέτοιο παράδειγμα αρχείου που παρέχει αυτήν την πληροφορία δίνεται στο [72] και παρατίθεται στο Σχήμα 41 για λόγους επίδειξης.

4.2.3 Στοιχειώδεις Υπηρεσίες του MPEG-M και Ενοποίησή τους

Οι στοιχειώδεις υπηρεσίες περιγράφονται στο τέταρτο τμήμα του προτύπου [90] και έχουν διαδραματίσει πολλαπλό ρόλο όχι μόνο στη χρήση, αλλά και στην εξέλιξη του προτύπου, καθώς αποτέλεσαν το σημείο αναφοράς για τη δημιουργία των μηχανών πρωτοκόλλου και τεχνολογίας που περιγράφηκαν στην προηγούμενη ενότητα. Στόχος αυτού του τέταρτου μέρους είναι να παράσχει τα πρωτόκολλα επικοινωνίας με υπηρεσίες, οι οποίες πραγματοποιούν θεμελιώδη λειτουργικότητα, όπως αυτή έχει προκύψει κατόπιν διαβουλεύσεων και απαιτήσεων που καταγράφηκαν λαμβάνοντας υπόψη διαφορετικές περιοχές όπου θα μπορούσε να εφαρμοστεί το μεσισμικό.

Αν και συμπεριλαμβάνονται σε ένα πρότυπο πολυμεσικής προέλευσης, οι λειτουργίες που επιτελούν οι στοιχειώδεις υπηρεσίες μπορούν να χαρακτηριστούν σε μεγάλο βαθμό ως γενικού σκοπού. Έτσι, το πρότυπο παρέχει πληροφορίες για υπηρεσίες που αφορούν το περιεχόμενο, από τη δημιουργία και την περιγραφή, μέχρι την προστασία και την παράδοσή του στον πελάτη, συσκευές, γεγονότα, άδειες χρήσης, υπηρεσίες, συμβόλαια, ακόμα και χρήστες. Η λειτουργικότητα που παρέχεται για καθέναν από τους προαναφερθέντες ρόλους δεν είναι πάντα η ίδια, αν και ακόμα και αυτή έχει ομαδοποιηθεί σε κατηγορίες, όπως η δημιουργία (π.χ. περιεχομένου, άδειας χρήσης, συμβολαίου κλπ), η αναζήτηση, η επεξεργασία και άλλα που περιγράφονται στο πρότυπο [90], ενώ ονομαστική λίστα είναι διαθέσιμη και στο [72].

Τα πρωτόκολλα αίτησης/απόκρισης της κάθε στοιχειώδους υπηρεσίας δίνονται σε μορφή XML μηνυμάτων (πρόκειται ουσιαστικά για τα μηνύματα που αναλαμβάνουν να

```

1  <?xml version="1.0" encoding="UTF-8"?>
2  <mxm:MXMConfiguration xmlns:mxm="org:iso:mpeg:mxm:configuration:schema"
3     xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
4     xsi:schemaLocation="org:iso:mpeg:mxm:configuration:schema mxmconfiguration.xsd">
5     <MXMParameters>
6         <entry key="username">JohnDoe</entry>
7         <ComplexParameter>
8             <bar:CustomMXMParam xmlns:bar="urn-x:bar">
9                 A complex MXM configuration parameter
10            </bar:CustomMXMParam>
11        </ComplexParameter>
12        <MXMEnginesFolder>/usr/bin/MXMEngines</MXMEnginesFolder>
13    </MXMParameters>
14    <mxm:MXMEngine id="0" type="MPEG21FileTE">
15        <ClassName>org.iso.mpeg.mxm.test.MPEG21FileEngine</ClassName>
16    </mxm:MXMEngine>
17    <mxm:MXMEngine id="1" type="MetadataTE" isDefault="true">
18        <ClassName>org.iso.mpeg.mxm.test.GenericMetadataEngine</ClassName>
19        <EngineParameters xsi:type="mxm:MXMGenericParameterType">
20            <entry key="verbosity">ALL</entry>
21        </EngineParameters>
22    </mxm:MXMEngine>
23    <mxm:MXMEngine id="2" type="MetadataTE">
24        <ClassName>org.iso.mpeg.mxm.test.MPEG7SMPMetadataEngine</ClassName>
25        <EngineParameters xsi:type="mxm:MXMGenericParameterType">
26            <entry key="verbosity">DEBUG</entry>
27        </EngineParameters>
28    </mxm:MXMEngine>
29    <mxm:MXMEngine id="3" type="DigitalItemTE">
30        <ClassName>org.iso.mpeg.mxm.test.MPEG7SMPMetadataEngine</ClassName>
31        <EngineParameters xsi:type="mxm:MXMGenericParameterType">
32            <entry key="verbosity">DEBUG</entry>
33        </EngineParameters>
34        <EngineDependencies>
35            <DependentMXMEngine id="1"/>
36        </EngineDependencies>
37    </mxm:MXMEngine>
38 </mxm:MXMConfiguration>

```

Σχήμα 41: Παράδειγμα Αρχείου Φορτώματος Υλοποιήσεων Μηχανών

διαχειριστούν οι μηχανές πρωτοκόλλου, όπως αναφέρθηκε παραπάνω στην Ενότητα 4.2.2), ενώ δίνεται και η αντίστοιχη λεκτική περιγραφή για καλύτερη κατανόηση. Επιπλέον αυτού, κάθε στοιχειώδης υπηρεσία συνοδεύεται και από μια αναπαράσταση της λειτουργίας της σε μορφή Business Processes Model and Notation (BPMN) [91].

Λόγω της στοιχειώδους φύσης τους, οι υπηρεσίες αυτές σπάνια χρησιμοποιούνται μόνες τους, παρά σε συνδυασμό με άλλες. Για παράδειγμα, για τη δημιουργία ενός ψηφιακού αντικείμενου χρησιμοποιείται η υπηρεσία CreateContent. Ωστόσο, για να προστεθεί αναγνωριστικό σε αυτό το ψηφιακό αντικείμενο, το οποίο είναι πολύ συνηθισμένη απαίτηση, χρειάζεται να χρησιμοποιηθεί στη συνέχεια η υπηρεσία IdentifyContent. Οδηγούμενη από αυτήν την παρατήρηση, η κοινότητα του MPEG-M οδηγήθηκε στην ενσωμάτωση ενός ακόμα, πέμπτου μέρους στο πρότυπο [92], το οποίο καθορίζει τον τρόπο με τον οποίο

θα γίνεται ο συνδυασμός των στοιχειωδών υπηρεσιών, σε μια ενιαία.

Αν και η αρχική προσέγγιση ήταν να χρησιμοποιηθεί το πρότυπο της BPMN για την αναπαράσταση μιας ενοποιημένης υπηρεσίας, τελικά αποφασίστηκε ότι το πρότυπο θα πρέπει να καθορίζει μόνο την πολιτική που θα ακολουθείται ως προς τη διαδικασία ενοποίησης. Ωστόσο, περιλαμβάνονται πληροφοριακές (informative) ενότητες, όπου δίνονται παραδείγματα ενοποιημένων υπηρεσιών με χρήση BPMN, η οποία έχει χρησιμοποιηθεί άλλωστε και στα πλαίσια της παρούσας διατριβής, όπως παρουσιάζεται στην επόμενη ενότητα.

Το πρωτόκολλο αίτησης/απόκρισης μιας ενοποιημένης υπηρεσίας θα πρέπει να ακολουθεί τη μορφή των στοιχειωδών υπηρεσιών, όπως δίνονται στο τέταρτο μέρος του προτύπου. Φυσικά, δε θα μπορούσε να καθοριστεί ένα γενικό πρωτόκολλο που να τις αφορά όλες, δεδομένου ότι οι υπηρεσίες μπορεί να συνδυαστούν με πολλούς τρόπους. Επιπλέον αυτού δε, στην ενοποιημένη υπηρεσία είναι δυνατό να περιλαμβάνονται και άλλες, μη πρότυπες υπηρεσίες ή ακόμα και άλλες ήδη υπάρχουσες ενοποιημένες υπηρεσίες (π.χ. αυτό μπορεί να συμβαίνει όταν χρησιμοποιείται μια ενοποιημένη υπηρεσία που παρέχει κάποια τρίτη οντότητα).

4.2.4 Υλοποίηση Υπηρεσίας Δημοσίευσης/Συνδρομής στο MPEG-M

Μια από τις συνεισφορές της παρούσας διατριβής στο πρότυπο είναι η εισαγωγή και υλοποίηση της μηχανής τεχνολογίας OverlayEngine [73] [75], η οποία παρέχει την προγραμματιστική διεπαφή για διατήρηση ενός υπερκείμενου δικτύου, καθώς και την αποστολή μηνυμάτων σε αυτό. Η υλοποίηση της εν λόγω μηχανής έχει γίνει με χρήση επιδημικού πρωτοκόλλου. Επιπλέον αυτής της μηχανής, προτάθηκε και έγινε αποδεκτή από την κοινότητα του MPEG-M η εισαγωγή της στοιχειώδους υπηρεσίας για πρωτόκολλα δημοσίευσης/συνδρομής PostContent [85], η οποία σε συνδυασμό με την OverlayEngine, καθώς και άλλες υπάρχουσες μηχανές και υπηρεσίες, μπορεί να υποστηρίξει πλήρως τα εν λόγω πρωτόκολλα.

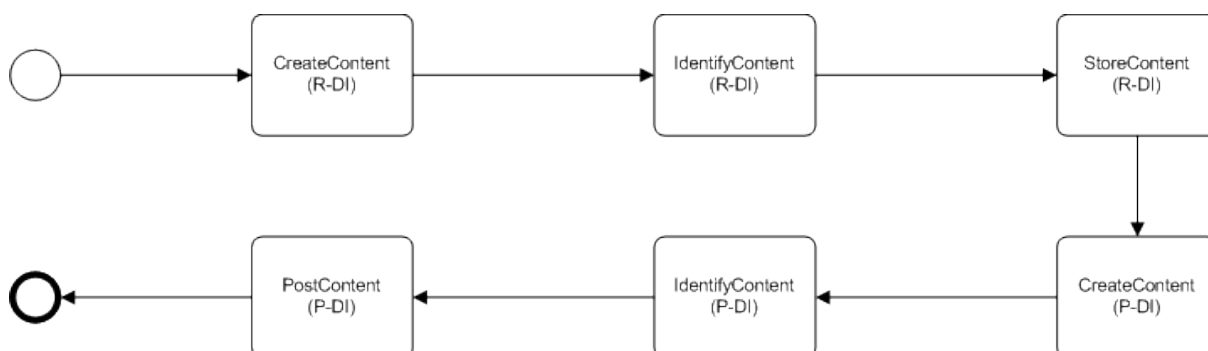
Αν και ο σχεδιασμός είναι τέτοιος, ώστε να μπορεί να υποστηριχθούν και οι τρεις ευρύτερες κατηγορίες πρωτοκόλλων δημοσίευσης/συνδρομής που περιγράφηκαν στην Ενότητα 3.1 (βασιζόμενα σε κάποιο θέμα, βασιζόμενα στο περιεχόμενο και βασιζόμενα στον τύπο), κεντρικό σημείο αποτέλεσαν τα πρωτόκολλα βασιζόμενα στο περιεχόμενο, καθώς είναι τα πιο γενικά. Ακολουθώντας τη συνολική λογική του MPEG-21 και του MPEG-M, όπου η πληροφορία είναι το επίκεντρο, η βασική ιδέα ήταν να δημιουργείται ένα ψηφιακό αντικείμενο για κάθε πληροφορία που πρόκειται να δημοσιευτεί.

Συγκεκριμένα το πρώτο βήμα της διαδικασίας δημοσίευσης είναι η δημιουργία ενός ψηφιακού αντικειμένου, το οποίο ενθυλακώνει όλες τις σχετικές με το δημοσιευόμενο περιεχόμενο πληροφορίες. Για το σκοπό αυτό χρησιμοποιούνται οι θεμελιώδεις υπηρεσίες CreateContent και IdentifyContent, με την πρώτη να το δημιουργεί και η δεύτερη να δίνει ένα

μοναδικό αναγνωριστικό στο εν λόγω ψηφιακό αντικείμενο. Δεδομένου ότι αφορά έναν πόρο (resource), το ψηφιακό αντικείμενο ονομάζεται ψηφιακό αντικείμενο πόρου (Resource Digital Item - R-DI). Στη συνέχεια, το R-DI αποθηκεύεται κάνοντας χρήση της στοιχειώδους υπηρεσίας StoreContent, ώστε να είναι διαθέσιμο από τους ενδιαφερόμενους συνδρομητές.

Καθώς το ψηφιακό αντικείμενο του πόρου συνήθως περιλαμβάνει μεγάλο αριθμό μεταδεδομένων, κάποια εκ των οποίων δεν είναι απαραίτητα για την ανακάλυψή της, προτάθηκε η δημιουργία ενός ψηφιακού αντικειμένου δημοσίευσης, το οποίο ενθυλακώνει μόνο την πληροφορία που είναι απαραίτητη σε ένα πρωτόκολλο δημοσίευσης/συνδρομής, το οποίο ονομάζεται ψηφιακό αντικείμενο δημοσίευσης (Publication Digital Item - P-DI). Ομοίως με το R-DI, και το P-DI δημιουργείται με χρήση των στοιχειωδών υπηρεσιών CreateContent και IdentifyContent.

Τελευταίο βήμα στη διαδικασία της δημοσίευσης αποτελεί η χρήση της στοιχειώδους υπηρεσίας PostContent, η οποία αναλαμβάνει να πάρει το ψηφιακό αντικείμενο δημοσίευσης και να το διοχετεύσει στο χρησιμοποιούμενο πρωτόκολλο, προκειμένου το τελευταίο να την καταστήσει διαθέσιμη στους συνδρομητές. Παράδειγμα τέτοιου πρωτοκόλλου αποτελεί το CCPS, που παρουσιάστηκε στο Κεφάλαιο 3.



Σχήμα 42: Ενοποιημένη Υπηρεσία Δημοσίευσης Περιεχομένου

Λαμβάνοντας υπόψη το πλήθος καθώς και την πολυπλοκότητα της παραπάνω διαδικασίας δημοσίευσης, το Σχήμα 42 παρουσιάζει το BPMN διάγραμμα μιας ενοποιημένης υπηρεσίας για δημοσίευση περιεχομένου. Τα επιμέρους στάδια που περιγράφηκαν νωρίτερα, δίνονται ως μια ακολουθία βημάτων μιας μεγαλύτερης ενοποιημένης υπηρεσίας, η οποία μπορεί να υποστηρίξει διαφανώς τη δημοσίευση περιεχομένου.

Η πορεία που ακολουθείται για τη συνδρομή σε περιεχόμενο είναι αντίστοιχη με αυτή της δημοσίευσης, εξαιρουμένων των βημάτων δημιουργίας του ψηφιακού αντικειμένου πόρου. Έτσι, στην προκειμένη περίπτωση, δημιουργείται ένα ψηφιακό αντικείμενο συνδρομής (Subscription Digital Item - S-DI), το οποίο δίνεται στη στοιχειώδη υπηρεσία PostContent, η οποία αναλαμβάνει να ξεκινήσει το πρωτόκολλο που χρησιμοποιείται για την εξεύρεση ταιριάζοντος δημοσιευμένου περιεχομένου. Το S-DI μπορεί να περιλαμβάνει μεταδεδομένα με λέξεις κλειδιά ή ακόμα και ένα ερώτημα (π.χ. για σημασιολογικά δεδομένα εκπεφρασμένο σε SPARQL Protocol and RDF Query Language - SPARQL). Πλέον αυτών, το

S-DI περιλαμβάνει και μια αίτηση για αναφορά γεγονότος (Event Report Request - ERR), η μορφή της οποίας καθορίζεται έτσι κι αλλιώς από το πρότυπο, έτσι ώστε όταν βρεθεί δημοσίευση που να αντιστοιχεί στη συνδρομή, να ειδοποιείται κατάλληλα ο συνδρομητής.

Κεφάλαιο 5

Το Κατανεμημένο Σύστημα Αρχείων EuterpeFS

5.1 Εισαγωγή

Ο διαμοιρασμός περιεχομένου αποτελεί διαχρονικό ζήτημα στον τομέα των κατανεμημένων συστημάτων. Αυτό αποδεικνύεται από τον όγκο της έρευνας που έχει διενεργηθεί όλα αυτά τα χρόνια σε διάφορες κατευθύνσεις, όπως τα ομότιμα δίκτυα, όπου διάφορες τεχνολογίες έχουν προταθεί, καθώς και άλλα πρωτόκολλα, τα οποία μπορεί να ακολουθούν είτε συμμετρική είτε κεντροκοποιημένη αρχιτεκτονική. Ωστόσο, αν και οι εν λόγω εργασίες έχουν δώσει μεγάλη ώθηση στη διάχυση του περιεχομένου, λειτουργούν σε ένα χαμηλότερο επίπεδο, που δε γίνεται αντιληπτό από τον τελικό χρήστη.

Η πιο κοινή αντίληψη των χρηστών για το περιεχόμενο αποτελεί το αρχείο. Διάφορα κατανεμημένα συστήματα αρχείων έχουν προταθεί στη βιβλιογραφία (με τα πιο βασικά από αυτά να επισημαίνονται στην επόμενη ενότητα), τα οποία επιδιώκουν να ικανοποιήσουν διαφορετικές απαιτήσεις, όπως διαθεσιμότητα, συνέπεια, αποδοτικότητα και ανοχή σε βλάβες. Ανεξάρτητα από το πεδίο όπου επικεντρώνονται όμως, όλα τα κατανεμημένα συστήματα αρχείων προσφέρουν μια διαφανή υπηρεσία αρχείων, έτσι ώστε οι χρήστες να προσπελαύνουν τα απομακρυσμένα αρχεία με τον ίδιο τρόπο όπως και τα τοπικά. Ωστόσο, ένα επίσης κοινό τους χαρακτηριστικό είναι ότι κάθε ένα κατανεμημένο σύστημα αρχείων στη βιβλιογραφία παρέχει τα δικές του τεχνολογίες, χωρίς να βασίζεται σε πρότυπα, ενώ γενικά αγνοούν και τον τρόπο αναπαράστασης που μπορεί να επιθυμεί ο εκάστοτε χρήστης, με βάση τα προσωπικά του χαρακτηριστικά και προτιμήσεις.

Επιπλέον, όσο αυξάνεται ο αριθμός των χρηστών, τόσο πιο πολύπλοκες γίνονται και οι απαιτήσεις για προστασία του περιεχομένου που διαθέτουν. Τα παραδοσιακά κατανεμημένα συστήματα αρχείων αδυνατούν να ανταποκριθούν στις απαιτήσεις ασφάλειας του σύγχρονου, ενοποιημένου περιβάλλοντος, όπου ένας πόρος μπορεί να διαμοιράζεται

μέσα από ποικίλες υπηρεσίες και πλατφόρμες, καθώς παρέχουν μηχανισμούς στενά συνδεδεμένους με την τεχνολογία τους. Από την άλλη, ένα τέτοιο πλαίσιο χρειάζεται ενιαίους μηχανισμούς προστασίας του περιεχομένου, ανεξάρτητα από την υπηρεσία διαμοιρασμού του.

Το κατανεμημένο σύστημα αρχείων EuterpeFS [93] παρέχει μια ενιαία, διαστρωματωμένη και βασισμένη σε πρότυπα λύση, η οποία παρέχει ένα σημασιολογικό μηχανισμό για την οργάνωση του κατανεμημένου περιεχομένου σε καταλόγους με βάση οντολογικές έννοιες, ενώ χρησιμοποιεί την προσέγγιση του MPEG για την προστασία του περιεχομένου με άδειες χρήσης. Έτσι, το EuterpeFS βασίζεται στα πρότυπα MPEG-21 και MPEG-M, τα οποία παρουσιάστηκαν στο Κεφάλαιο 4. Συγκεκριμένα, ενθυλακώνει μηχανισμούς όπως το ψηφιακό αντικείμενο και η γλώσσα περιγραφής δικαιωμάτων, ενώ βασίζεται και στα πρωτόκολλα του MPEG-M για να μπορέσει να χρησιμοποιήσει τους εν λόγω μηχανισμούς.

Μία ακόμα ιδιότητα του EuterpeFS είναι ότι λειτουργεί σε υψηλότερο επίπεδο από τα συνήθη κατανεμημένα συστήματα αρχείων. Έτσι, ένα αρχείο στο EuterpeFS δεν είναι παρά ένας δείκτης στο πραγματικό αρχείο (πόρο), μαζί με το πρωτόκολλο που θα πρέπει να χρησιμοποιηθεί για να προσπελαστεί ο εν λόγω πόρος. Έτσι, μπορεί ο πόρος αυτός να είναι μια ιστοσελίδα, οπότε χρησιμοποιείται το HTTP για την προσπέλαση, ένα βίντεο, οπότε χρησιμοποιείται ένα πρωτόκολλο ροής βίντεο, ένα απομακρυσμένο αρχείο, το οποίο μπορεί να διαμοιράζεται για παράδειγμα με BitTorrent είτε ακόμα και να παρέχεται από ένα άλλο κατανεμημένο σύστημα αρχείων (π.χ. NFS). Με αυτόν τον τρόπο, οι χρήστες είναι σε θέση να χρησιμοποιούν τις ήδη υπάρχουσες τεχνολογίες που καλύπτουν τις απαιτήσεις τους σε αποδοτικότητα, μέσα από το EuterpeFS.

5.2 Απαιτήσεις Κατανεμημένων Συστημάτων Αρχείων

Οι γενικές απαιτήσεις που διέπουν το σχεδιασμό των κατανεμημένων συστημάτων αρχείων είναι ουσιαστικά όμοιες με αυτές που πηγάζουν από το σχεδιασμό ενός οποιουδήποτε κατανεμημένου συστήματος. Στην παρούσα Ενότητα δίνονται αυτές οι απαιτήσεις όπως έχουν αρχικά παρουσιαστεί στο [94].

5.2.1 Διαφάνεια

Για να θεωρείται ένα κατανεμημένο σύστημα διαφανές, θα πρέπει να καλύπτει ένα σύνολο επιμέρους ιδιοτήτων, οι οποίες παρουσιάζονται στη συνέχεια. Σε κάθε περίπτωση, είναι γεγονός ότι η διαφάνεια θεωρείται θεμελιώδης ιδιότητα των κατανεμημένων συστημάτων και δει των κατανεμημένων συστημάτων αρχείων, όπου ο χρήστης θα πρέπει να έχει παρόμοια εμπειρία χρήσης με αυτήν ενός οποιουδήποτε, συμβατικού, τοπικού συστήματος αρχείων. Συγκεκριμένα, ένα κατανεμημένο σύστημα αρχείων θα πρέπει να ικανοποιεί τις παρακάτω κατηγορίες διαφάνειας.

5.2.1.1 Διαφάνεια Πρόσβασης

Η πρόσβαση στα αρχεία θα πρέπει να γίνεται με ομοιόμορφο τρόπο, είτε μιλάμε για τοπικό είτε για κατανεμημένο σύστημα αρχείων. Έτσι, ο χρήστης θα πρέπει να είναι σε θέση να χρησιμοποιήσει το ίδιο σύνολο εντολών/οδηγιών για να μπορέσει να πραγματοποιήσει λειτουργίες εισόδου/εξόδου πάνω στα αρχεία του. Συνεπώς, το κατανεμημένο σύστημα αρχείων θα πρέπει να υλοποιεί, αν όχι όλες, ένα μεγάλο αριθμό από λειτουργίες που είναι διαθέσιμες από κάποιο τοπικό σύστημα αρχείων.

5.2.1.2 Διαφάνεια Τοποθεσίας

Ο χώρος των ονομάτων ενός κατανεμημένου συστήματος αρχείων δε θα πρέπει να είναι διαφορετικός από αυτόν ενός τοπικού συστήματος αρχείων, δίνοντας έτσι τη δυνατότητα στις εφαρμογές να μπορούν να προσπελαίνουν αρχεία διαφορετικών συστημάτων με πανομοιότυπο τρόπο. Για παράδειγμα, η διαφάνεια τοποθεσίας μας εξασφαλίζει ότι μια εφαρμογή θα μπορούσε να έχει πρόσβαση κάνοντας χρήση του ίδιου συνόλου εντολών στους παρακάτω καταλόγους:

- /usr/foo/local: κατάλογος του τοπικού συστήματος αρχείων
- /usr/foo/remote: αναρτημένος κατάλογος του κατανεμημένου συστήματος αρχείων.

5.2.1.3 Διαφάνεια Μετακίνησης

Στο σύγχρονο κινητό περιβάλλον, θα πρέπει να θεωρείται δεδομένο ότι οι χρήστες που διαμοιράζονται τα αρχεία τους μπορεί να βρίσκονται σε κίνηση. Έτσι, υπάρχει μεγάλη πιθανότητα, ένα αρχείο, το οποίο παρέχεται από κάποιον εξυπηρετητή που απαντά σε μια συγκεκριμένη διεύθυνση, να αλλάξει διεύθυνση -ή ακόμα και να αλλάξει εξυπηρετητή και να παρέχεται από κάποιον άλλο. Ένα κατανεμημένο σύστημα αρχείων θα πρέπει να λαμβάνει υπόψη του αυτά τα ενδεχόμενα και να είναι σε θέση να παρέχει τις δύο παραπάνω κατηγορίες διαφάνειας (πρόσβασης και τοποθεσίας) ακόμα και αν τα αρχεία αλλάζουν θέση και δεν παρέχονται πάντα από τον ίδιο εξυπηρετητή.

5.2.1.4 Διαφάνεια Επίδοσης

Ο αναμενόμενος φόρτος στον οποίο μπορεί να ανταπεξέλθει ένα κατανεμημένο σύστημα αρχείων προσδιορίζεται μέσα σε κάποιο εύρος τιμών, το οποίο και δίνεται από το σχεδιαστή του. Η διαφάνεια επίδοσης εξασφαλίζει ότι όσο αυτός ο φόρτος βρίσκεται μέσα στο εύρος που έχει δοθεί ως σχεδιαστική παράμετρος του συστήματος, τότε και η απόδοσή του θα παραμένει, σχετικά, σταθερή.

5.2.1.5 Διαφάνεια Κλιμάκωσης

Η κλιμάκωση είναι μια ακόμα *de facto* απαίτηση ενός κατανεμημένου συστήματος και αναφέρεται στη δυνατότητα του τελευταίου να μπορεί να ανταποκρίνεται καλά ακόμα και όταν αυξάνονται οι αιτήσεις εξυπηρέτησης. Έτσι, για παράδειγμα, ακόμα και για την περίπτωση ενός δημοφιλούς αρχείου, το κατανεμημένο σύστημα αρχείων θα πρέπει να είναι σε θέση να υποστηρίξει πολλαπλές παράλληλες αιτήσεις πρόσβασης σε αυτό.

5.2.2 Παράλληλη Ενημέρωση Αρχείων

Σε ένα κατανεμημένο σύστημα αρχείων υπάρχουν πολλοί χρήστες, οι οποίοι μπορεί να αλληλεπιδρούν με τα ίδια αρχεία. Έτσι, είναι συχνή η περίπτωση όπου ένας χρήστης έχει ανοικτό και διαβάζει κάποιο αρχείο, το οποίο, εκείνη ακριβώς τη στιγμή επεξεργάζεται κάποιος άλλος χρήστης που το τροποποιεί και το αποθηκεύει. Επιπλέον, μπορεί και ο πρώτος χρήστης να αποφασίσει να κάνει κάποια αλλαγή στο αρχείο αφού έχει αποθηκευτεί τροποποιημένο από τον προηγούμενο χρήστη. Είναι ουσιαστική απαίτηση των κατανεμημένων συστημάτων αρχείων να υποστηρίζουν αυτές τις περιπτώσεις και να διατηρούν συνεπείς εκδόσεις των αρχείων ακόμα και όταν έχουν πολλοί χρήστες πρόσβαση σε αυτά. Οι μηχανισμοί που χρησιμοποιούνται στην πράξη είναι αντίστοιχοι με αυτούς που χρησιμοποιούνται και για να ελεγχθεί η ταυτόχρονη πρόσβαση σε ένα τοπικό αρχείο από πολλούς χρήστες -χρησιμοποιώντας τεχνολογίες όπως αρχεία λουκέτα ή κλειδωμά εγγραφής (άνοιγμα μόνο για ανάγνωση), διατήρηση προσωρινών αντιγράφων (swap files) και άλλα.

5.2.3 Αντίγραφα Αρχείων

Η τεχνική των αντιγράφων χρησιμοποιείται ευρέως στα κατανεμημένα συστήματα για δύο βασικούς λόγους: ανοχή βλαβών και αποδοτικότητα. Αν και δεν είναι αυστηρή προϋπόθεση για ένα κατανεμημένο σύστημα αρχείων να παρέχει εγγυήσεις αντιγράφων, πρόκειται για μια πολύ διαδεδομένη τεχνική, αφού με κόστος χώρου (ο οποίος θεωρείται “φθηνός” σε σχέση με το εύρος ζώνης το οποίο θα αυξάνεται εφόσον θα πρέπει να γίνονται συχνές αιτήσεις σε ένα εξυπηρετητή για ένα αρχείο) μπορούμε και εξισορροπούμε το φόρτο στους εξυπηρετητές συστημάτων αρχείων. Με αυτό τον τρόπο, μπορούμε να επιλέξουμε να εξυπηρετούμε τους χρήστες από διαφορετικούς εξυπηρετητές με βάση διάφορα κριτήρια, όπως η τοποθεσία των χρηστών, ή ο φόρτος σε κάποιον εξυπηρετητή, οι οποίοι, όμως, θα πρέπει να διατηρούν πανομοιότυπα αντίγραφα των αρχείων.

Το τελευταίο αυτό σημείο σημαίνει ότι, ένα κατανεμημένο σύστημα αρχείων θα πρέπει να διατηρεί *συνεπή* αντίγραφα των αρχείων. Συνεπώς, αν κάποιος χρήστης τροποποιήσει κάποιο αρχείο που προσπελαύνει μέσω κάποιου εξυπηρετητή Α, ενώ το ίδιο αρχείο παρέχεται σε άλλους χρήστες από κάποιον άλλο εξυπηρετητή Β, τότε το ίδιο το σύστημα αρχείων θα πρέπει να ενημερώσει άμεσα τον εξυπηρετητή Β για τη νέα έκδοση του αρ-

χείου.

5.2.4 Ετερογένεια

Η ετερογένεια είναι ένας ακόμα παράγοντας, ο οποίος έχει γίνει ιδιαίτερα σημαντικό στις μέρες μας. Για να πληρεί αυτήν την απαίτηση, ένα κατανεμημένο σύστημα αρχείων θα πρέπει να παρέχει υλοποιήσεις για διαφορετικές πλατφόρμες υλικού και λογισμικού (π.χ. λειτουργικών συστημάτων), ενσωματωμένο, όμως, κάθε φορά στο εικονικό σύστημα αρχείων του εκάστοτε υπολογιστικού συστήματος. Έτσι, ένα κατανεμημένο σύστημα αρχείων, θα πρέπει να υλοποιεί το UNIX Virtual File System (VFS) για να μπορεί να χρησιμοποιηθεί σε ένα UNIX σύστημα αλλά και το αντίστοιχο εικονικό σύστημα αρχείων των Windows για να μπορούν να έχουν πρόσβαση σε αυτό χρήστες και αυτού του λειτουργικού συστήματος και με αυτόν τον τρόπο να μπορούν να διαμοιράζονται τα αρχεία τους.

5.2.5 Ανοχή Βλαβών

Η ανοχή βλαβών αποτελεί ένα από τα σημαντικότερα κεφάλαια των σύγχρονων κατανεμημένων συστημάτων και, μάλιστα, πολλές φορές από τα δυσκολότερα. Τα κατανεμημένα συστήματα αρχείων, εντούτοις, δε θεωρούνται τόσο ευαίσθητα σε αυτόν τομέα, καθώς χρησιμοποιώντας σημειολογία *at most once* (ή *at least once* με φροντίδα για τις πολλαπλές ενημερώσεις που μπορεί να λάβουν χώρα σε αυτήν την περίπτωση), μη κρατώντας κατάσταση για τις εφαρμογές πελάτη που μπορεί να κάνουν χρήση της υπηρεσίας συστήματος αρχείων και κάνοντας χρήση αντιγράφων, καταφέρνουν να παρουσιάζουν μεγάλη ανοχή σε βλάβες.

5.2.6 Συνέπεια

Η συνέπεια μπορεί να θεωρηθεί ο συμπληρωματικός παράγοντας των αντιγράφων, κυρίως γιατί τα τελευταία έχουν αρνητική επίδραση πάνω της. Συγκεκριμένα, η συνέπεια αναφέρεται στην ικανότητα που έχει ένα κατανεμημένο σύστημα αρχείων να κρατά ενήμερα όλα τα αντίγραφα ενός αρχείου που μπορεί να βρίσκονται διασκορπισμένα στο δίκτυο, με τη μικρότερη δυνατή καθυστέρηση. Κάτι τέτοιο, βέβαια, στην πράξη μπορεί να αποδειχθεί ιδιαίτερα δύσκολο, καθώς η συνέπεια δεν αναφέρεται μόνο σε αντίγραφα που μπορεί να διατηρούνται στους εξυπηρετητές, αλλά και στους πελάτες των κατανεμημένων συστημάτων αρχείων, και τα οποία μπορεί να τροποποιούνται ανά πάσα στιγμή από διαφορετικούς χρήστες. Μάλιστα, ένας ακόμα παράγοντας που καθιστά τη συνέπεια μια δύσκολη απαίτηση είναι ότι για να την αυξήσουμε, θα πρέπει να αυξήσουμε το κόστος επικοινωνίας, το οποίο ειδικά σε ένα σύγχρονο κινητό περιβάλλον είναι έτσι κι αλλιώς ακριβό.

5.2.7 Ασφάλεια

Τα τοπικά συστήματα αρχείων παρέχουν μηχανισμούς ελέγχου πρόσβασης στα δεδομένα χρησιμοποιώντας λίστες ελέγχου πρόσβασης. Έτσι, για παράδειγμα, στο UNIX, κάθε αρχείο έχει έναν ιδιοκτήτη χρήστη και μια ιδιοκτήτρια ομάδα χρηστών, ενώ μπορούμε να ορίσουμε και δικαιώματα ανάγνωσης, εγγραφής και εκτέλεσης σε αυτούς τους χρήστες και τις ομάδες. Με μια πρώτη ματιά και με βάση τη διαφάνεια πρόσβασης και τοποθεσίας που παρέχει ένα κατανεμημένο σύστημα αρχείων, θα μπορούσαν να εφαρμοστούν οι ίδιοι μηχανισμοί ασφάλειας με τα τοπικά συστήματα αρχείων.

Η βασική διαφορά έγκειται στο γεγονός ότι ο χρήστης του τοπικού συστήματος αρχείων έχει ήδη αυθεντικοποιηθεί στο τοπικό σύστημα και, με βάση αυτόν τον κωδικό, έχει αποκτήσει πρόσβαση στα δεδομένα πάνω στα οποία έχει δικαιώματα. Έτσι, σε πλήρη αντιστοιχία, και ένα κατανεμημένο σύστημα αρχείων θα πρέπει να είναι σε θέση να αυθεντικοποιήσει πρώτα το χρήστη που αιτείται πρόσβασης στα αρχεία και στη συνέχεια να αποφασίσει για το αν θα δώσει την πρόσβαση. Επιπλέον, ένα κατανεμημένο σύστημα αρχείων χρειάζεται έναν ακόμα μηχανισμό ασφάλειας για τη μεταφορά των αρχείων πάνω από το δίκτυο (π.χ. κρυπτογράφηση).

5.2.8 Αποδοτικότητα

Η αποδοτικότητα είναι ένας παράγοντας που μετράται με διαφορετικές μονάδες, ανάλογα με την περιοχή εφαρμογής του κατανεμημένου συστήματος αρχείων. Έτσι, για παράδειγμα, για άλλες περιπτώσεις μπορεί να είναι ο χρόνος προσπέλασης, ενώ για άλλες οι περιορισμένες ανάγκες για τοπικό αποθηκευτικό χώρο. Σε κάθε περίπτωση, όμως, θα πρέπει να παρέχεται ένα σύστημα αρχείων που να μπορεί να χρησιμοποιηθεί από τις εφαρμογές κάθε πλατφόρμας χωρίς να θέτει περιορισμούς και να επηρεάζει αρνητικά τη συνολική απόδοση του συστήματος.

5.3 Αρχιτεκτονική Κατανεμημένων Συστημάτων Αρχείων

Η σχεδίαση και υλοποίηση ενός κατανεμημένου συστήματος αρχείων εναπόκειται κάθε φορά στον αντίστοιχο προγραμματιστή/αναλυτή. Η βάση αυτού του σχεδιασμού είναι η κάλυψη των απαιτήσεων που παρουσιάστηκαν στην παραπάνω Ενότητα 5.2. Εντούτοις, στη βιβλιογραφία [94] [95] συναντάμε τα βασικά χαρακτηριστικά που θα πρέπει να έχουν τα θεμελιώδη τμήματα ενός κατανεμημένου συστήματος αρχείων.

5.3.1 Υπηρεσία Αρχείων

Η υπηρεσία αρχείων, όχι μόνο σε ένα κατανεμημένο αλλά σε οποιοδήποτε σύστημα αρχείων, είναι υπεύθυνη να δίνει πρόσβαση στις εφαρμογές στα αρχεία, έχοντας γνώση του περιβάλλοντος διεπαφής που χρησιμοποιεί κάθε λειτουργικό σύστημα (π.χ. για το UNIX το εικονικό σύστημα αρχείων - VFS). Αυτή η διεπαφή δίνει τη δυνατότητα για πρόσβαση τόσο στα δεδομένα όσο και σε ιδιότητες των αρχείων, οι οποίες είναι εξίσου απαραίτητες για την απεικόνιση και την πρόσβαση σε αυτά.

Η θεμελιώδης δομή που χρησιμοποιείται στο UNIX για να συμπεριλάβει τα βασικά χαρακτηριστικά ενός αρχείου είναι η *stat* [96]. Αυτή η δομή περιλαμβάνει πληροφορίες απαραίτητες για το λειτουργικό σύστημα ώστε να αποκτήσει πρόσβαση στα δεδομένα που ενθυλακώνει το αρχείο και βρίσκονται εγγεγραμμένα στο σκληρό δίσκο του υπολογιστή. Έτσι, ένα σύστημα αρχείων θα πρέπει να είναι σε θέση να αναγιγνώσκει αλλά και να εγγράφει/τροποποιεί αυτή τη δομή με βάση τις αλλαγές στα χαρακτηριστικά των αρχείων. Συνεπώς, με δεδομένο ότι αυτή η διαδικασία γίνεται στην πλευρά του χρήστη, καθώς στο δικό του υπολογιστικό σύστημα θα γίνει η προσπέλαση των αρχείων, ένα κατανεμημένο σύστημα θα πρέπει να εξασφαλίζει με διαφάνεια την πρόσβαση στα χαρακτηριστικά των αρχείων, με τις εφαρμογές να κάνουν χρήση του περιβάλλοντος διεπαφής του VFS, το οποίο με τη σειρά του υλοποιείται αναλόγως από το κατανεμημένο σύστημα αρχείων.

5.3.1.1 Χαρακτηριστικά Αρχείων

Στην Ενότητα αυτή παρουσιάζονται τα χαρακτηριστικά των αρχείων, όπως αυτά ορίζονται στην προαναφερθείσα δομή *stat*. Σε αυτό το σημείο θα πρέπει να σημειωθεί ότι αυτά τα χαρακτηριστικά δεν αποτελούν μέρος των δεδομένων που φέρει το αρχείο, αλλά της πληροφορίας που φέρει το ίδιο το σύστημα αρχείων για την πρόσβαση στα αρχεία. Επιπλέον, τόσο ένα τοπικό όσο και ένα κατανεμημένο σύστημα αρχείων θα πρέπει να έχει πρόσβαση σε αυτά τα χαρακτηριστικά -η υλοποίηση είναι αυτή που αλλάζει κατά περίπτωση. Έτσι, η δομή *stat* περιλαμβάνει την εξής πληροφορία:

- Το αναγνωριστικό της συσκευής που περιλαμβάνει το αρχείο. Αυτό το αναγνωριστικό προκύπτει από τη σύνθεση του μείζονος (*major*) και του ελάσσονος (*minor*) αναγνωριστικού της συσκευής, όπως αυτά έχουν ανατεθεί από τον πυρήνα του συστήματος. Κάνοντας χρήση αυτής της πληροφορίας, η υπηρεσία αρχείων είναι σε θέση να εντοπίσει σε ποια συσκευή βρίσκονται αποθηκευμένα τα δεδομένα που ενθυλακώνει το αρχείο. Σε ένα κατανεμημένο σύστημα αρχείων αυτό το αναγνωριστικό μπορεί να είναι διαφορετικό σε κάθε πελάτη (π.χ. αυτό συμβαίνει και στο Network File System - NFS) ανάλογα με την υλοποίηση.
- Τον αριθμό ευρετηρίου (*i-node*) του αρχείου στη συσκευή όπου είναι αποθηκευμένο. Αυτό το πεδίο χρησιμοποιείται για να είναι σε θέση το σύστημα αρχείων να γνωρίζει

από ποια διεύθυνση του αποθηκευτικού μέσου θα ξεκινήσει να διαβάζει τα bytes του αρχείου. Σε αντιστοιχία με το αναγνωριστικό της συσκευής, έτσι και ο αριθμός ευρετηρίου μπορεί να αντιστοιχεί στον πραγματικό αριθμό ευρετηρίου στον εξυπηρετητή αρχείων, ώστε να διευκολύνεται η αναζήτηση στον εξυπηρετητή ή να αντιστοιχεί σε μια τοπική έκδοση του αρχείου, ώστε να μπορεί το σύστημα αρχείων να προσπελάσει το τοπικό αντίγραφο του αρχείου. Και σε αυτήν την περίπτωση, η επιλογή γίνεται με βάση το σχεδιασμό του κάθε κατανεμημένου συστήματος αρχείων.

- Τον τρόπο πρόσβασης στο αρχείο. Το πεδίο αυτό περιγράφει τόσο τον τύπο του αρχείου όσο και το είδος της πρόσβασης που μπορούν να έχουν στο αρχείο ο ιδιοκτήτης χρήστης του (u), η ιδιοκτήτρια ομάδα χρηστών (g) και οι υπόλοιποι (o). Ο τύπος του αρχείου μπορεί να είναι: κανονικό αρχείο, κατάλογος αρχείων, συσκευή χαρακτηριστων, συσκευή μπλοκ, σωλήνας (pipe), σύνδεσμος και στόμιο. Το είδος της πρόσβασης μπορεί να είναι: ανάγνωση (r), εγγραφή (w), εκτέλεση (x). Έτσι, για κάθε ένα από τα παραπάνω υποκείμενα (u, g, o) διατίθενται 3 bits που υποδεικνύουν αντίστοιχα αν έχουν δικαίωμα ανάγνωσης, εγγραφής και εκτέλεσης. Σε ένα κατανεμημένο σύστημα αρχείων απαιτείται ιδιαίτερη προσοχή σε αυτήν την ιδιότητα, καθώς είναι έτσι κι αλλιώς δύσκολο να προσδιοριστεί ο ιδιοκτήτης ή η ομάδα ιδιοτήτων του αρχείου και, συνεπώς, είναι δύσκολο να οριστούν ή να τροποποιηθούν από την πλευρά του πελάτη αυτές οι τιμές.
- Τον αριθμό των σκληρών συνδέσμων (hard links) του αρχείου. Όταν υπάρχει σκληρός σύνδεσμος ανάμεσα σε δύο αρχεία, τότε τα αρχεία αυτά έχουν ίδιο αριθμό ευρετηρίου.
- Το αναγνωριστικό χρήστη του ιδιοκτήτη του αρχείου. Κάθε χρήστης του UNIX έχει ένα μοναδικό αναγνωριστικό χρήστη, το οποίο και του ανατίθεται κατά τη δημιουργία του λογαριασμού του στο σύστημα. Ένα κατανεμημένο σύστημα αρχείων θα πρέπει σε αυτήν την περίπτωση να λάβει υπόψη του το αναγνωριστικό χρήστη, όπως αυτό δίνεται από την πλευρά του πελάτη του συστήματος αρχείων. Συνεπώς, θα πρέπει είτε να υπάρχει αντιστοιχία ανάμεσα στα ονόματα χρηστών του τοπικού συστήματος (πλευρά πελάτη) με το απομακρυσμένο (πλευρά εξυπηρετητή) είτε να υπάρχει κάποιο άλλο σχήμα αυθεντικοποίησης του χρήστη πελάτη στον εξυπηρετητή, το οποίο όμως θα πρέπει να γίνεται με διαφάνεια, χρησιμοποιώντας δηλαδή την ίδια διεπαφή για τον έλεγχο πρόσβασης.
- Το αναγνωριστικό ιδιοκτήτριας ομάδας του αρχείου. Όπως και οι χρήστες, έτσι και οι ομάδες χρηστών έχουν μοναδικό αναγνωριστικό στο σύστημα -όταν εγγράφεται ένας χρήστης, τότε αντιστοιχίζεται και σε ομάδες χρηστών (π.χ. η τυπική ομάδα χρηστών του UNIX στην οποία υπάγονται οι χρήστες είναι η 'users'). Σε ένα κατανεμημένο σύστημα αρχείων, θα πρέπει να ληφθεί φροντίδα για τον έλεγχο πρόσβασης των ομάδων χρηστών, ανάλογη με αυτήν για τους χρήστες.

- Το αναγνωριστικό ειδικής συσκευής που μπορεί να αντιπροσωπεύει αυτό το αρχείο. Στο UNIX ένα αρχείο μπορεί να χρησιμοποιηθεί όχι μόνο για πρόσβαση σε δεδομένα αποθηκευμένα σε ένα αποθηκευτικό μέσο αλλά και για πρόσβαση σε συσκευές ή διεργασίες. Συνήθως τα κατανεμημένα συστήματα δε χρειάζεται να έχουν ξεχωριστή υλοποίηση για το χειρισμό τέτοιων αρχείων, καθώς κάτι τέτοιο είναι εκτός των πλαισίων λειτουργίας τους, και χρησιμοποιούν υλοποιήσεις τοπικών συστημάτων αρχείων.
- Το μέγεθος του αρχείου σε bytes. Αυτή η ιδιότητα είναι απαραίτητη για να μπορέσει η υπηρεσία πελάτη να αναγνώσει το αρχείο από τη συσκευή αποθήκευσης, καθώς θα πρέπει να γνωρίζει πόσα bytes θα πρέπει να διαβάσει και να επιστρέψει στο χρήστη. Σε ένα κατανεμημένο σύστημα αρχείων, αυτή η λειτουργία γίνεται στην πλευρά του εξυπηρετητή και επιστρέφεται το μέγεθος του αρχείου στην πλευρά του πελάτη, ο οποίος αιτείται την ανάγνωση του αρχείου.
- Το μέγεθος μπλοκ που θα χρησιμοποιηθεί για τις λειτουργίες εισόδου/εξόδου. Το μέγεθος αυτό μπορεί να ποικίλλει με βάση το είδος των αρχείων για τα οποία χρησιμοποιείται το σύστημα αρχείων. Έτσι, ένα μεγάλο μέγεθος μπλοκ μπορεί να σημαίνει μεγάλους χρόνους αναμονής για να διαβαστεί ή να εγγραφεί ένα κομμάτι του αρχείου, ενώ από την άλλη ένα μικρό μέγεθος μπλοκ μπορεί να σημαίνει πολλαπλές κλήσεις συστήματος για εντολές εισόδου/εξόδου. Σε ένα κατανεμημένο σύστημα αρχείων αυτό το μέγεθος μπλοκ αντιστοιχίζεται στον αριθμό των bytes που διαβάζει (γράφει) η υπηρεσία αρχείων από τον (στον) εξυπηρετητή αρχείων.
- Τον αριθμό των μπλοκ που καταλαμβάνει το αρχείο στη συσκευή αποθήκευσης. Αυτός ο αριθμός μπορεί να ποικίλλει, ανάλογα με την υλοποίηση, στα 512 bytes ή στο μέγεθος μπλοκ που έχει δοθεί στην προηγούμενη μεταβλητή.
- Τη χρονική στιγμή της τελευταίας προσπέλασης των δεδομένων του αρχείου. Σε ένα κατανεμημένο σύστημα αρχείων αυτή η τιμή είναι άμεσα συνδεδεμένη με τη συνέπεια, καθώς η τελευταία είναι η ιδιότητα που καθορίζει το πόσο συχνά θα ανανεώνεται η πληροφορία που κρατείται στον εξυπηρετητή για τις μεταβολές που υφίσταται το αρχείο. Επιπλέον, καθώς αυτός ο χρόνος έχει να κάνει απλά με την ανάγνωση και όχι με την τροποποίηση, μπορεί να μην κρίνεται σκόπιμη η συχνή ανανέωσή του, καθώς το κόστος επικοινωνίας μπορεί να είναι υψηλό για την ακρίβεια μιας τέτοιας μεταβλητής.
- Τη χρονική στιγμή της τελευταίας τροποποίησης των δεδομένων του αρχείου. Σε αντίστοιχία με την προηγούμενη ιδιότητα, και αυτή σχετίζεται άμεσα με τη συνέπεια. Η διαφορά είναι ότι σε αυτήν την περίπτωση, η ανάγκη για ανανέωση είναι πιο σημαντική, καθώς μπορεί να χρησιμοποιηθεί για να γνωρίζουν οι πελάτες αν έχουν “φρέσκο” αντίγραφο του αρχείου (δηλαδή αν ο χρόνος τελευταίας τροποποίησης σε κάθε πελάτη είναι ίδιος με αυτόν στον εξυπηρετητή).

- Τη χρονική στιγμή τελευταίας τροποποίησης της κατάστασης του αρχείου, δηλαδή κάποιας από τις ιδιότητες του αρχείου και όχι των δεδομένων. Έτσι, αυτός ο χρόνος ανανεώνεται όταν τροποποιείται ο τρόπος πρόσβασης του αρχείου, ο ιδιοκτήτης του, κλπ.

5.3.1.2 Μοντέλο Προσπέλασης Αρχείων

Το μοντέλο προσπέλασης αρχείων αναφέρεται στον τρόπο με τον οποίο η υπηρεσία αρχείων προσπελαύνει τα απομακρυσμένα αρχεία. Έτσι, τα πιο διαδεδομένα μοντέλα προσπέλασης αρχείων [95] είναι το μοντέλο μεταφοράς αρχείων και το μοντέλο απομακρυσμένης πρόσβασης. Η διαφορά τους έγκειται στην προσέγγιση που ακολουθούν όταν ο χρήστης ζητάει να ανοίξει ένα αρχείο. Έτσι, στο μεν μοντέλο μεταφοράς αρχείων, όταν ο χρήστης ζητάει ένα αρχείο, τότε το αρχείο αυτό αντιγράφεται τοπικά στο υπολογιστικό του σύστημα, στο δε μοντέλο απομακρυσμένης πρόσβασης, οι αιτήσεις πραγματοποιούνται απευθείας στην πλευρά του εξυπηρετητή κατά απαίτηση του χρήστη.

Το μοντέλο μεταφοράς αρχείων έχει το βασικό πλεονέκτημα της απλής υλοποίησης. Έτσι, η βασική του λειτουργικότητα έγκειται στην πραγματοποίηση μιας μεταφοράς του απομακρυσμένου αρχείου και στη συνέχεια, μπορεί να αναλάβει το τοπικό σύστημα αρχείων να εξυπηρετεί το μεγαλύτερο μέρος των εντολών πάνω σε αυτό το αρχείο. Αυτή η προσέγγιση, επίσης, μπορεί να είναι ιδιαίτερα αποδοτική σε περιβάλλοντα χαμηλής συνδεσιμότητας, όπου η συσκευή του πελάτη μπορεί να μην έχει συνέχεια επικοινωνία με τον εξυπηρετητή. Σε αυτήν την περίπτωση, το να διατηρείται ένα τοπικό αντίγραφο του αρχείου δίνει τη δυνατότητα στην εφαρμογή πελάτη να προσπελαύνει το αρχείο οποιαδήποτε στιγμή. Τα μειονεκτήματα αυτής της προσέγγισης αφορούν τη συνέπεια και το κόστος επικοινωνίας. Συγκεκριμένα, η συνέπεια αναφέρεται στο γεγονός ότι με αυτήν την προσέγγιση, ουσιαστικά αναγόμεστε στο πρόβλημα διατήρησης συνεπών αντιγράφων ενός αρχείου σε διάφορα σημεία. Έτσι, το κατανεμημένο σύστημα αρχείων θα πρέπει να είναι σε θέση να ενημερώνει τους πελάτες της υπηρεσίας αρχείων σχετικά με τροποποιήσεις που γίνονται σε κάποιο ανοικτό αρχείο από άλλους πελάτες. Επιπλέον, ο εξυπηρετητής θα πρέπει να δύναται να διατηρεί την πιο φρέσκια έκδοση αρχείων, ακόμα και μετά από ταυτόχρονες τροποποιήσεις ή, τουλάχιστον, ο χρήστης να γνωρίζει αν τελικά οι αλλαγές του αποθηκεύτηκαν χωρίς να υπάρχει κάποιο πρόβλημα λόγω κάποιας άλλης αντίστοιχης ενέργειας που έλαβε χώρα παράλληλα. Το κόστος επικοινωνίας μπορεί να είναι μεγάλο σε αυτήν την περίπτωση, αν και το μέγεθος των αρχείων είναι μεγάλο -πολλά δε μάλλον αν ο χρήστης χρειάζεται μόνο ένα μικρό μέρος πολλών αρχείων. Τέλος, ειδικά στην περίπτωση μεγάλων αρχείων, αυτό το μοντέλο μπορεί να επιβαρύνει σημαντικά τον πελάτη, καθώς χρειάζεται χώρο στο αποθηκευτικό του μέσο για να μπορέσει να διατηρήσει τα ανοικτά αρχεία.

Το μοντέλο απομακρυσμένης πρόσβασης έχει, κατά κανόνα, πιο πολύπλοκη υλο-

ποίηση καθώς η πλευρά του πελάτη θα πρέπει να πραγματοποιεί απομακρυσμένες κλήσεις στον εξυπηρετητή, ο οποίος θα πραγματοποιεί τις πραγματικές κλήσεις πάνω στα αρχεία. Έτσι, σε μια απλουστευμένη μορφή, θα μπορούσαμε να ισχυριστούμε ότι το εν λόγω μοντέλο μεταφέρει την εντολή του πελάτη στον εξυπηρετητή, ο οποίος την εκτελεί και επιστρέφει τα αποτελέσματα στον πελάτη. Το πλεονέκτημα αυτής της προσέγγισης είναι ότι δεν απαιτεί πολύ χώρο στην πλευρά του πελάτη, αφού κάθε φορά μεταφέρεται μόνο το κομμάτι του αρχείου που αιτείται ο τελευταίος. Από την άλλη, οι συχνές πράξεις προσπέλασης του αρχείου από τον πελάτη αυξάνουν το κόστος επικοινωνίας.

Η επιλογή του μοντέλου προσπέλασης των αρχείων τελικά εξαρτάται από το πεδίο εφαρμογής του κατανεμημένου συστήματος αρχείων. Έτσι, οι σχεδιαστές λαμβάνουν υπόψη τους το πλαίσιο λειτουργίας των συστημάτων αρχείων και στη συνέχεια επιλέγουν το ένα μοντέλο ή το άλλο. Επίσης, υβριδικά μοντέλα μπορεί να χρησιμοποιηθούν ώστε να βελτιστοποιείται η απόδοση. Για παράδειγμα, θα μπορούσε να χρησιμοποιηθεί το μοντέλο απομακρυσμένης πρόσβασης σε συνδυασμό με τη διατήρηση κρυφής μνήμης (cache) στην πλευρά του πελάτη, ώστε να μη χρειάζονται επαναλαμβανόμενες κλήσεις πάνω από το δίκτυο για κομμάτια του αρχείου τα οποία έχουν ήδη μεταφερθεί σε προηγούμενη χρονική στιγμή. Περισσότερες λεπτομέρειες σχετικά με τις υλοποιήσεις κατανεμημένων συστημάτων αρχείων δίνονται στο Κεφάλαιο 5.4.

5.3.2 Υπηρεσία Καταλόγων

Η βασική προτεραιότητα της υπηρεσίας καταλόγων είναι η ονομασία των αρχείων και η διατήρηση μιας κατά απαίτηση δομής καταλόγων ανά πελάτη, χωρίς να επηρεάζεται ο εξυπηρετητής αρχείων. Για να πετύχει αυτό το στόχο, χρειάζεται να συνεργαστεί με την υπηρεσία αρχείων, όπως θα αναλυθεί παρακάτω. Σε ένα σύγχρονο σύστημα αρχείων, η πλευρά του πελάτη θα πρέπει να είναι σε θέση να οργανώνει τοπικά τους προσαρτημένους καταλόγους, εξασφαλίζοντας πάντα στην πρόσβαση στον εξυπηρετητή. Ακόμα περισσότερο, η υπηρεσία καταλόγου θα πρέπει να εξασφαλίζει διαφάνεια τοποθεσίας, όπου ακόμα και όταν μετακινείται ο εξυπηρετητής, δε θα πρέπει να επηρεάζονται οι πελάτες.

Για να μπορέσουν να ανταποκριθούν σε αυτές τις απαιτήσεις, τα κατανεμημένα συστήματα αρχείων πλέον χρησιμοποιούν ένα επίπεδο χώρο ονομάτων αντιστοιχίζοντας στα αρχεία που βρίσκονται στον εξυπηρετητή μοναδικά αναγνωριστικά (Unique File Identifiers - UFIDs) μέσω των οποίων μπορούν οι πελάτες να τα προσπελούν, ανεξάρτητα από την οργάνωση στους τοπικούς καταλόγους.

Έτσι, σε περίπτωση που θέλουμε να δημιουργήσουμε ένα νέο αρχείο, η υπηρεσία αρχείων θα πρέπει να παράξει ένα UFID το οποίο θα δοθεί στην υπηρεσία καταλόγων, η οποία με τη σειρά της είναι υπεύθυνη για την αντιστοίχιση του UFID με το πραγματικό όνομα αρχείου και περιεχόμενου καταλόγου, ώστε ο κάθε πελάτης να μπορεί να προσπελάσει το νέο αρχείο μέσα στο μονοπάτι όπου έχει οριστεί κατά τη δημιουργία του.

5.3.3 Προσάρτηση Απομακρυσμένων Καταλόγων

Εικονικό Σύστημα Αρχείων Το εικονικό σύστημα αρχείων (Virtual File System - VFS) αποτελεί το υψηλότερο επίπεδο αφαίρεσης ενός συστήματος αρχείων και παρέχει τη απαραίτητη διαφάνεια στις εφαρμογές να χρησιμοποιούν το ίδιο περιβάλλον διεπαφής, χωρίς να έχουν γνώση της υποκείμενης υλοποίησης του εκάστοτε συστήματος αρχείων. Για να μπορέσει να ανταπεξέλθει σε αυτές τις απαιτήσεις, το VFS διατηρεί ξεχωριστές δομές [97], οι οποίες περιγράφονται σε αυτήν την Ενότητα.

Οι δύο βασικές δομές που χρησιμοποιεί το VFS για να λειτουργήσει σωστά είναι το υπερμπλόκ (superblock) και το ευρετήριο κόμβου (i-node) κάθε αρχείου. Στα UNIX συστήματα, το υπερμπλόκ είναι μια δεσμευμένη περιοχή του κάθε διαμερίσματος, η οποία ξεκινά αμέσως μετά το μπλοκ εκκίνησης (boot block), είναι δηλαδή το δεύτερο σε σειρά μπλοκ και έχει ίδιο μέγεθος με το πρώτο, δηλαδή 1024 bytes [98]. Το υπερμπλόκ στην ουσία περιγράφει το πώς είναι διαμορφωμένο το διαμέρισμα στο οποίο ανήκει, έχει ένα δείκτη στο πρώτο i-node και περιλαμβάνει πληροφορία για το μέγεθος του μπλοκ καθώς και άλλες πληροφορίες συγκεκριμένες ανά σύστημα αρχείων. Το i-node έχει ως βασικό σκοπό να δείχνει πού βρίσκονται στο αποθηκευτικό μέσο τα δεδομένα που περιλαμβάνει το αρχείο που αντιπροσωπεύει.

Όταν προσαρτάται κάποιος κατάλογος, είτε είναι τοπικό διαμέρισμα είτε κάποιο απομακρυσμένο, το VFS επιχειρεί να αναγνωρίσει το σύστημα αρχείων που υλοποιεί την προσπέλαση αυτού του καταλόγου. Αν δεν αναγνωριστεί ο κατάλογος, τότε το VFS περνάει τον έλεγχο στον πυρήνα, ο οποίος με τη σειρά του είτε θα επιστρέψει μήνυμα λάθους είτε θα επιχειρήσει να βρει κάποιο πρόγραμμα που θα του επιτρέψει να χρησιμοποιήσει το σύστημα αρχείων που ζητήθηκε κατά την προσάρτηση.

Έτσι, στην περίπτωση ενός κατανεμημένου συστήματος αρχείων, το VFS θα κληθεί να το αναγνωρίσει όταν επιχειρηθεί προσάρτηση κάποιου απομακρυσμένου καταλόγου στο τοπικό σύστημα. Εφόσον γίνει αυτό το βήμα, τότε οι εφαρμογές μπορούν να συνεχίσουν να επικαλούνται το περιβάλλον διεπαφής του VFS και οι κλήσεις που πραγματοποιούν σε κάθε προσαρτημένο κατάλογο να εκτελούνται από το αντίστοιχο σύστημα αρχείων.

5.3.3.1 Σύστημα Αρχείων στο Χώρο Χρηστών

Ο χώρος χρηστών στα μονολιθικά λειτουργικά συστήματα, όπως το LINUX, χρησιμεύει γιατί επιτρέπει στους χρήστες να φορτώσουν δικό τους κώδικα στον πυρήνα, χωρίς να χρειάζεται να πειράξουν κάποιο πρόγραμμα στην περιοχή του πυρήνα. Σημειώνεται ότι, παραδοσιακά, τα συστήματα αρχείων για το LINUX ήταν γραμμένα στον πυρήνα.

Το σύστημα αρχείων στο χώρο χρηστών (File system in USErspace - FUSE) δίνει τη δυνατότητα στους χρήστες να γράψουν τα δικά τους συστήματα αρχείων, χωρίς να χρειάζε-

ται να προγραμματίσουν τον πυρήνα. Συγκεκριμένα, το FUSE παρέχει ένα περιβάλλον διεπαφής καθώς και τις αντίστοιχες βιβλιοθήκες, με τις οποίες οι χρήστες-προγραμματιστές μπορούν να υπερβούν τις κλήσεις συστήματος των συστημάτων αρχείων με δικές τους υλοποιήσεις.

Λόγω της ευελιξίας που παρέχει, το FUSE έχει γίνει ιδιαίτερα δημοφιλές για την ανάπτυξη συστημάτων αρχείων, και κυρίως κατανεμημένων όπου οι προγραμματιστές δίνουν περισσότερο βάση στην επικοινωνία παρά στο τρόπο πρόσβασης στην πληροφορία στο τοπικό επίπεδο, και έτσι με το FUSE είναι σε θέση να συνδυάζουν δικές τους υλοποιήσεις με τις αντίστοιχες του εκάστοτε υποκείμενου συστήματος αρχείων. Στην επίσημη ιστοσελίδα του FUSE [99] διατηρείται λίστα με τις υπάρχουσες υλοποιήσεις συστημάτων αρχείων που βασίστηκαν σε αυτό.

5.4 Μελέτες Περίπτωσης Κατανεμημένων Συστημάτων Αρχείων

5.4.1 Sun Network File System

Το Sun Network File System (NFS) [100] [101] βασίζεται στο πρωτόκολλο NFS [102], το οποίο και δίνει τη δυνατότητα σε οποιαδήποτε εφαρμογή πελάτη να πραγματοποιεί απομακρυσμένες κλήσεις συστήματος αρχείων. Το NFS υποστηρίζει ένα μεγάλο εύρος λειτουργικών συστημάτων, από το MS-DOS μέχρι το VMS [103] αν και στην παρούσα Ενότητα, μένοντας στο πνεύμα και της διατριβής θα παρουσιαστεί η προσέγγιση υλοποίησης για το UNIX.

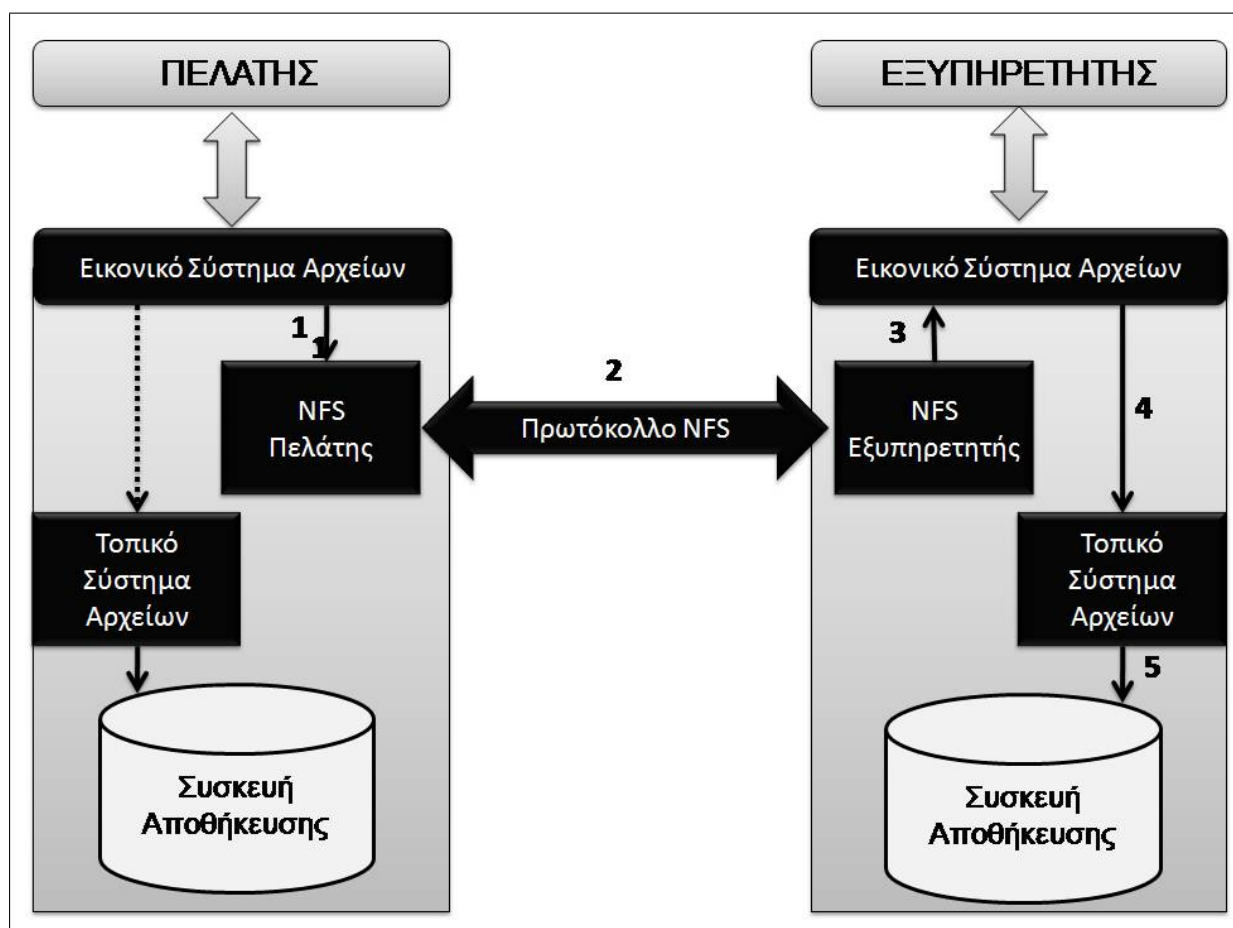
Το NFS ακολουθεί παραδοσιακή αρχιτεκτονική πελάτη-εξυπηρετητή, οι οποίοι επικοινωνούν κάνοντας χρήση του αντίστοιχου πρωτοκόλλου, το οποίο με τη σειρά του βασίζεται στο μοντέλο επικοινωνίας της απομακρυσμένης κλήσης διεργασίας (Remote Procedure Call - RPC). Μάλιστα, το σύστημα RPC της, η Sun το ανέπτυξε για τους σκοπούς του NFS πρωτοκόλλου [94]. Επιπλέον, το NFS μπορεί να χρησιμοποιεί είτε TCP είτε UDP πρωτόκολλο για την επικοινωνία πελάτη-εξυπηρετητή.

Στο NFS κάθε υπολογιστικό σύστημα που πρόκειται να μοιράζει τα αρχεία του θα πρέπει να τρέχει τον εξυπηρετητή NFS. Ο ρόλος του τελευταίου είναι να επικοινωνεί από το ένα άκρο με τον πελάτη NFS μέσω του αντίστοιχου πρωτοκόλλου NFS και από το άλλο να μεταφράζει τις αιτήσεις του πελάτη, σε δικές του αιτήσεις στο τοπικό του εικονικό σύστημα αρχείων. Έτσι, στην ουσία, ο εξυπηρετητής δρα ως πληρεξούσιος του πελάτη στην πλευρά του υπολογιστικού συστήματος που έχει τα αρχεία και αναλαμβάνει να πραγματοποιεί τις κλήσεις για λογαριασμό του. Ο εξυπηρετητής του NFS βρίσκεται στο χώρο του πυρήνα του λειτουργικού συστήματος για να μπορεί να έχει πρόσβαση στο VFS του συστήματος που μοιράζει τα αρχεία.

Από την πλευρά του, ο πελάτης NFS υλοποιεί τις κλήσεις συστήματος του VFS και

είναι ενσωματωμένος στον πυρήνα του λειτουργικού συστήματος του πελάτη. Όταν γίνεται κάποια κλήση προς αρχείο προσαρτημένο στο NFS, ο πελάτης αναλαμβάνει να μεταφέρει αυτήν την κλήση στην πλευρά του εξυπηρετητή, κάνοντας χρήση του πρωτοκόλλου NFS.

Στο Σχήμα 43 φαίνεται η αρχιτεκτονική του κατανεμημένου συστήματος αρχείων NFS. Όπως φαίνεται και από τα αριθμημένα βήματα, όταν γίνει μια κλήση για ένα αρχείο που εξυπηρετείται από το NFS στην πλευρά του πελάτη, τότε το VFS θα δρομολογήσει αυτήν την αίτηση στον πελάτη NFS, στη συνέχεια εκείνος θα χρησιμοποιήσει το πρωτόκολλο NFS για να στείλει την αίτηση στον εξυπηρετητή NFS, ο οποίος με τη σειρά του θα κάνει αυτήν την αίτηση στο τοπικό του VFS. Το VFS στην πλευρά του εξυπηρετητή, τότε, θα αναγνωρίσει το σύστημα αρχείων για το οποίο προορίζεται η αίτηση και θα το χρησιμοποιήσει για να την πραγματοποιήσει. Το αποτέλεσμα θα δοθεί από το σύστημα αρχείων στον εξυπηρετητή πίσω στο VFS και θα ακολουθηθεί η ίδια πορεία βημάτων προς τα πίσω μέχρι να φτάσουμε στην πλευρά του πελάτη και το εκεί VFS.



Σχήμα 43: Η Αρχιτεκτονική του NFS

Το NFS ακολουθεί το μοντέλο απομακρυσμένης πρόσβασης των αρχείων, πράγμα που σημαίνει ότι θα πρέπει να πραγματοποιεί απομακρυσμένες κλήσεις για κάθε κλήση

που πραγματοποιείται στην πλευρά του πελάτη. Όπως αναφέρθηκε, όμως, και στο τέλος της Ενότητας 5.3.1.2, το NFS χρησιμοποιεί ένα υβριδικό σχήμα, σύμφωνα με το οποίο διατηρείται κρυφή μνήμη τόσο στην πλευρά του πελάτη όσο και στην πλευρά του εξυπηρετητή.

Ο εξυπηρετητής χρησιμοποιεί την κλασική προσέγγιση των κρυφών μνημών που ακολουθείται από οποιοδήποτε τοπικό σύστημα αρχείων, κρατώντας για κάποιο χρονικό διάστημα στην κρυφή μνήμη τα αντίγραφα των αρχείων που διαβάζονται, οπότε τις σχετικές αιτήσεις των πελατών τις εξυπηρετεί από την κρυφή μνήμη όταν μπορεί. Το πρόβλημα, όμως, που υπάρχει σε ένα κατανεμημένο σύστημα αρχείων είναι αυτό της συνέπειας των αντιγράφων. Έτσι, ο εξυπηρετητής θα πρέπει να πάρει απόφαση τι θα κάνει εφόσον του έρθει εντολή να γράψει κάποιο αρχείο -αν δηλαδή θα πρέπει να το γράψει κατευθείαν στο αποθηκευτικό μέσο (write through) ή να περιμένει πρώτα τελική επιβεβαίωση από τον πελάτη (write commit). Μέχρι και την έκδοση 3, αυτό ήταν μια παράμετρος του εξυπηρετητή. Από την έκδοση 4, όμως, το NFS δίνει τη δυνατότητα στον πελάτη να αποφασίσει το σχήμα που θα ακολουθηθεί για την κρυφή μνήμη. Έτσι, μπορεί για παράδειγμα, να δοθεί η δυνατότητα στον πελάτη να κλειδώσει κάποιο αρχείο, το οποίο χρησιμοποιεί ή στο οποίο σκοπεύει να γράψει, ώστε να μην μπορεί να το τροποποιήσει κάποιος άλλος πελάτης στο μεταξύ και να υπάρχει συνέπεια των αντιγράφων στους διάφορους πελάτες.

5.4.2 Andrew File System

Το Andrew File System (AFS) [104] ήρθε στην ουσία για να καλύψει την εν γένει αδυναμία του NFS σε περιβάλλοντα χαμηλής συνδεσιμότητας -ιδιότητα που εκμεταλλεύτηκε πολύ περισσότερο ο απόγονός του, το κατανεμημένο σύστημα αρχείων Coda [105].

Η διεργασία εξυπηρετητή στο AFS αποκαλείται Vice, ενώ η διεργασία πελάτη Venus. Το AFS, σε αντίθεση με το NFS, ακολουθεί το μοντέλο μεταφοράς αρχείων και έτσι η κάθε διεργασία πελάτη διατηρεί τη δική της έκδοση των αρχείων τοπικά, για κάθε αρχείο το οποίο αιτείται από τον εξυπηρετητή. Μάλιστα, αυτά τα αρχεία τα οποία μεταφέρονται στον πελάτη, διατηρούνται εκεί για μεγάλο χρονικό διάστημα -ακόμα και μετά από επανεκκίνηση. Μπορούμε, λοιπόν, να θεωρήσουμε ότι το AFS κάνει την υπόθεση μιας μεγάλης κρυφής μνήμης, την οποία αναλαμβάνει να διατηρεί ενήμερη.

Η συνέπεια της κρυφής μνήμης είναι μια από τις σημαντικότερες απαιτήσεις που πρέπει να καλύπτει το AFS. Για να το επιτύχει αυτό, βασίζεται κυρίως στην πλευρά του εξυπηρετητή Vice, ο οποίος είναι υπεύθυνος να ενημερώνει τους πελάτες που έχουν αντίγραφα των αρχείων για τυχόν αλλαγές σε αυτά. Έτσι, όταν ένας πελάτης Venus ανοίξει ένα αρχείο (δηλαδή το αντιγράψει από τον εξυπηρετητή στην τοπική του κρυφή μνήμη), τότε ο εξυπηρετητής κρατά μια μεταβλητή κατάστασης, στην οποία κρατά το γεγονός ότι ο εν λόγω πελάτης έχει ανοικτό το σχετικό αρχείο. Όταν κάποιος άλλος πελάτης τροποποιήσει το αρχείο αυτό, τότε ο Vice αναλαμβάνει να ελέγξει αυτή τη μεταβλητή για πελάτες που μπορεί να έχουν ανοικτό αυτό το αρχείο και να τους ενημερώσει ότι το αντίγραφό τους

είναι πλέον άκυρο και θα πρέπει να ζητήσουν πλέον καινούριο, αν το επιθυμούν.

Για να βελτιώσει την απόδοσή του, το AFS κάνει κάποιες υποθέσεις σχετικά με το μέγεθος των αρχείων και το πλήθος των αναφορών σε αυτά, οι σημαντικότερες των οποίων συνοψίζονται στο [94]:

- Τα αρχεία είναι μικρά με μέγεθος μικρότερο των 10KB το καθένα. Αυτό όπως αναφέρθηκε άλλωστε και στην Ενότητα 5.3.1.2 είναι και ένα βασικό μειονέκτημα του μοντέλου μεταφοράς αρχείων που υποστηρίζει το AFS.
- Οι αιτήσεις των πελατών για ανάγνωση των αρχείων είναι πολύ συχνότερες, σε σχέση με τις αιτήσεις για εγγραφή (περίπου 6 φορές περισσότερες).
- Τα αρχεία προσπελούνται κυρίως σειριακά και όχι τυχαία.
- Τα περισσότερα αρχεία διαβάζονται και γράφονται από ένα μόνο χρήστη -ακόμα και στην περίπτωση που έχουμε πολλούς χρήστες, συνήθως μόνο ένας είναι αυτός που κάνει τροποποιήσεις στο αρχείο.
- Τα αρχεία προσπελούνται κατά ριπάς -ένα αρχείο που προσπελαύθηκε πρόσφατα, είναι πολύ πιθανό να ξαναζητηθεί σύντομα.

5.4.3 Lustre

Το Lustre είναι ένα κατανεμημένο σύστημα αρχείων σχεδιασμένο για συστάδες υπολογιστικών συστημάτων και αναπτύχθηκε στα πλαίσια του προγράμματος DARPA High Productivity Computing Systems (HPCS). Λόγω της καλής του κλιμάκωσης, το Lustre συναντά ευρεία εφαρμογή σε ένα πλατύ φάσμα εταιρειών που πρέπει να διατηρούν και να επεξεργάζονται πολύ μεγάλους όγκους δεδομένων, όπως πετρελαϊκές, κατασκευαστικές, πολυμέσων ή ακόμα και παρόχους υπηρεσιών διαδικτύου [106].

Η αρχιτεκτονική του Lustre θεωρεί δύο θεμελιώδη τμήματα: τους Metadata Servers (MDS) και τους Object Storage Servers (OSS) [107]. Οι πρώτοι, ο αριθμός των οποίων γενικά μπορεί να εκτείνεται από μερικές δεκάδες σε εκατοντάδες, χρησιμοποιούνται για να αποθηκεύουν τα μεταδεδομένα που σχετίζονται με τα αρχεία. Οι δεύτεροι, ο αριθμός των οποίων μπορεί να εκτείνεται από μερικές εκατοντάδες σε χιλιάδες, χρησιμοποιούνται για να αποθηκεύουν αυτή καθαυτή την πληροφορία που φέρουν τα αρχεία.

Πιο συγκεκριμένα, ένας MDS φέρει ένα δικό του, ιδιωτικό σύστημα αρχείων, το οποίο χρησιμοποιεί για να αποθηκεύει τα μεταδεδομένα, το οποίο ονομάζεται Metadata Target (MDT). Τα μεταδεδομένα που αποθηκεύονται στους MDS έχουν να κάνουν τόσο με τα ονόματα των αρχείων και των καταλόγων, τα οποία επηρεάζονται και από τις αντίστοιχες λειτουργίες του συστήματος αρχείων για δημιουργία ή μετονομασία αρχείων ή συνδέσμων όσο και με τα δικαιώματα πρόσβασης στα αρχεία. Το υλικό που υποστηρίζει

τους MDS είναι υψηλής απόδοσης για να μπορεί να υποστηρίξει τις μεγάλες απαιτήσεις των χρηστών του.

Οι Object Storage Servers (OSS) φέρουν και αυτοί το δικό τους ιδιωτικό σύστημα αρχείων, το οποίο ονομάζεται Object Storage Target (OST), για να αποθηκεύουν τα δεδομένα των αρχείων. Έτσι, οι OSS αναλαμβάνουν τις λειτουργίες εισόδου εξόδου των αρχείων, ενώ υποστηρίζουν την κατανομή κομματιών των αρχείων σε πολλά σημεία (OSS) ώστε να υπάρχει ομοιομορφία στην κατανομή του περιεχομένου για όλους τους χρήστες.

Ακολουθώντας αυτήν την τακτική, το Lustre καταφέρνει να κάνει καλύτερη χρήση των διαθέσιμων πόρων του, οι οποίοι θα πρέπει να είναι έτσι κι αλλιώς πολλοί για να υποστηρίξουν τις απαιτητικές εφαρμογές, εκμεταλλευόμενο το υποκείμενο σύστημα σε όλο του το εύρος. Μάλιστα, για να μην υπάρχουν καθυστερήσεις μεταξύ της ανάγνωσης των μεταδεδωμένων και των πραγματικών δεδομένων, αλλά ακόμα και μεταξύ διαδοχικών αναγνώσεων των δεδομένων (αφού τα τελευταία βρίσκονται αποθηκευμένα σε διαφορετικούς OSS) το Lustre θεωρεί ένα δίκτυο πολύ υψηλών ταχυτήτων που ενώνει τα διάφορα συστατικά του.

Όταν ένας χρήστης επιχειρήσει να ανοίξει ένα αρχείο του Lustre, τότε ουσιαστικά μεταφέρεται από τον MDS που διατηρεί τα μεταδεδωμένα του αρχείου μια αντιστοίχιση του ονόματος του αρχείου με τις τοποθεσίες (OST) όπου διατηρούνται τα πραγματικά δεδομένα του αρχείου. Με αυτόν τον τρόπο, και ανάλογα με την προσπέλαση, η εφαρμογή του πελάτη υπολογίζει κάθε φορά με βάση αυτήν την αντιστοίχιση, από ποιον OST θα ζητήσει τα δεδομένα που επιχειρεί να προσπελάσει ο χρήστης. Συνεπώς, ο MDS δε συμμετέχει σε κανένα στάδιο της διαδικασίας εισόδου εξόδου, φέροντας μόνο την πληροφορία που χρειάζεται για να πραγματοποιήσει τις κατάλληλες αιτήσεις ο πελάτης.

5.5 Ψηφιακά Αντικείμενα για Αναπαράσταση Αρχείων

Το EuterpeFS χρησιμοποιεί τη δομή του ψηφιακού αντικειμένου για την αναπαράσταση των αρχείων. Έτσι, δίνεται η δυνατότητα να υπάρχει ένα πλαίσιο το οποίο μπορεί να αναπαραστήσει οποιονδήποτε πόρο και με το EuterpeFS αυτόν τον πόρο να μπορεί να τον προσπελαύνει ο χρήστης σαν αρχείο. Για παράδειγμα, ένα τέτοιο αρχείο θα μπορούσε να είναι από ένα απλό κείμενο, μέχρι μια ιστοσελίδα, μια υπηρεσία, μια ροή δεδομένων κλπ. Έτσι, όταν ο χρήστης επιχειρήσει να ανοίξει ένα αρχείο, το οποίο αναφέρεται σε μια ιστοσελίδα, τότε θα ανοίξει ο φυλλομετρητής του πηγαίνοντάς τον στην αναφερόμενη ιστοσελίδα.

Το ψηφιακό αντικείμενο που αναπαριστά ένα αρχείο, το οποίο είναι σε θέση να επεξεργαστεί το EuterpeFS θα έχει τη μορφή που φαίνεται στο Σχήμα 44. Η δομή του ψηφιακού αντικειμένου στο σύνολο, καθώς και των επιμέρους στοιχείων που ενθυλακώνει έχει περιγραφεί στην Ενότητα 4.1. Η σειρά με την οποία δίνονται τα επιμέρους συστατικά του είναι

```

1 <DIDL xmlns="urn:mpeg:mpeg21:2002:02-DIDL-NS"
2   xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
3   xmlns:dii="urn:mpeg:mpeg21:2002:01-DII-NS"
4   xsi:schemaLocation="urn:mpeg:mpeg21:2002:02-DIDL-NS didl.xsd">
5
6 <Container>
7   <Descriptor id="identifier">
8     <Statement>
9       <dii:Identifier>...</dii:Identifier>
10    </Statement>
11  </Descriptor>
12
13  <Descriptor id="metadata">
14    <Statement>...</Statement>
15  </Descriptor>
16
17  <Descriptor id="license">
18    <Statement>
19      <licenseGroup>
20        <license>...</license>
21      </licenseGroup>
22    </Statement>
23  </Descriptor>
24
25  <Item>
26    <Descriptor id="stat">
27      <Statement>
28        <size>...</size>
29      </Statement>
30    </Descriptor>
31
32    <Component>
33      <Resource ref="...">
34    </Component>
35  </Item>
36
37 </Container>
38 </DIDL>

```

Σχήμα 44: Μορφή Ψηφιακού Αντικειμένου για Αναπαράσταση Αρχείου στο EuterpeFS

ενδεικτική και δεν είναι απαραίτητο να είναι πάντα η ίδια. Εντούτοις, όπως θα περιγραφεί και παρακάτω, κάθε ένα από αυτά τα συστατικά επιτελεί συγκεκριμένες λειτουργίες και ανάλογα με το ποια περιλαμβάνονται, παρέχεται και η αντίστοιχη λειτουργικότητα από το EuterpeFS.

Το πρώτο στοιχείο του ψηφιακού αντικειμένου αρχείου αφορά το μοναδικό αναγνωριστικό του. Η ύπαρξή του είναι απαραίτητη, καθώς το EuterpeFS το χρησιμοποιεί για να αναπαραστήσει με μοναδικό τρόπο το αντίστοιχο αρχείο στο δίκτυο, υιοθετώντας τη λογική του NFS όπου κάθε αρχείο φέρει ένα UUID. Επιπλέον, αν δε δίνεται όνομα του αρχείου σε επόμενο στοιχείο, το αναγνωριστικό αυτό έχει και το ρόλο του ονόματος στην τοπική συσκευή του πελάτη.

Στη συνέχεια περιλαμβάνονται τα μεταδεδομένα που περιγράφουν τον πόρο, ο

οποίος αναπαρίσταται από το ψηφιακό αντικείμενο. Τα μεταδεδομένα αυτά μπορούν να έχουν οποιαδήποτε μορφή, από απλό κείμενο μέχρι σημασιολογικά. Ωστόσο, ένα μεγάλο μέρος της λειτουργικότητας του EuterpeFS βασίζεται σε αυτήν την πληροφορία. Συγκεκριμένα, προτείνεται η χρήση σημασιολογικών τεχνολογιών (π.χ. οντολογία), ώστε να γίνεται στη συνέχεια από την υπηρεσία αρχείων η κατάλληλη αντιστοίχιση σε καταλόγους.

Επόμενο περιγραφικό στοιχείο του ψηφιακού αντικειμένου είναι η άδεια χρήσης που συνοδεύει τον πόρο που αναπαριστά. Αυτή η άδεια χρήσης εκφράζεται στην πρότυπη γλώσσα έκφρασης δικαιωμάτων του MPEG-21 (REL) και μπορεί είτε να μην υπάρχει καν είτε να μην περικλείεται στο ψηφιακό αντικείμενο. Ειδικά για τη δεύτερη περίπτωση, το πρότυπο παρέχει τη δυνατότητα σε ένα ψηφιακό αντικείμενο να αναφέρει μια άδεια χρήσης, η οποία είναι διαθέσιμη από μια απομακρυσμένη τοποθεσία. Σε κάθε περίπτωση, όμως, η εφαρμογή πελάτη είναι σε θέση να τη λάβει και να τη χρησιμοποιήσει προκειμένου να πραγματοποιήσει τον έλεγχο πρόσβασης που έχει ζητήσει μέσω αυτής ο πάροχος του εν λόγω πόρου.

Τέλος, το ψηφιακό αντικείμενο αρχείου θα πρέπει να περιλαμβάνει πληροφορίες σχετικές με την πρόσβαση στον πόρο, καθώς και δεδομένα που χρειάζεται η δομή stat, προκειμένου να μπορέσει να αναπαραστήσει το αρχείο στο σκληρό δίσκο. Έτσι, για το πρώτο, δίνεται η διαδρομή προς τον πόρο με τη μορφή ενός ενιαίου εντοπιστή τοποθεσίας (Uniform Resource Locator - URL), ενώ για το δεύτερο δίνονται σε μορφή XML τα στοιχεία που χρειάζεται η δομή stat. Απαραίτητο για την τελευταία είναι το μέγεθος του αρχείου, το οποίο δίνεται σε bytes.

5.6 Έλεγχος Πρόσβασης στο EuterpeFS

Το κατανεμημένο σύστημα αρχείων EuterpeFS βασίζεται στη γλώσσα περιγραφής δικαιωμάτων του MPEG-21 προκειμένου ο πάροχος του περιεχομένου να δώσει τους κανόνες ελέγχου πρόσβασης με τη μορφή μιας άδειας χρήσης, η οποία ενσωματώνεται στο ψηφιακό αντικείμενο αρχείου. Με αυτόν τον τρόπο, όχι μόνο δίνεται στον πάροχο η δυνατότητα χρήσης μιας πρότυπης γλώσσας, σχεδιασμένης ειδικά για αυτό το σκοπό, αλλά επιπλέον το ίδιο ψηφιακό αντικείμενο μπορεί να χρησιμοποιηθεί ομοίως προστατευμένο σε οποιοδήποτε περιβάλλον, δεδομένου ότι στηρίζεται σε πρότυπη τεχνολογία.

Ο έλεγχος και, κατά επέκταση, η εξουσιοδότηση για την πρόσβαση σε κάποιον πόρο γίνεται με χρήση στοιχειωδών υπηρεσιών του MPEG-M. Έτσι, από τη μία πλευρά υποστηρίζεται η δημιουργία κανόνων που αφορούν μεμονωμένους χρήστες ή ακόμα και ομάδες χρηστών, οι οποίοι όμως ομαδοποιούνται από κάποια χαρακτηριστικά, τα οποία και πάλι πρέπει να αποδείξουν ότι κατέχουν νόμιμα. Ειδικά το κομμάτι που αφορά τους κατόχους χαρακτηριστικών (στην ορολογία του MPEG-21) αποτελεί μια από τις πλέον ισχυρές βάσεις υποστήριξης ετερογένειας, δεδομένου ότι ένας χρήστης αποκτά ένα χαρακτηριστικό μέσω

μιας ειδικής άδειας, η οποία (στην τυπική περίπτωση) είναι ψηφιακά υπογεγραμμένη. Επιπλέον, αυτά τα χαρακτηριστικά δύνανται να αποκτηθούν κατά οποιονδήποτε τρόπο, αρκεί να περικλείονται στην πρότυπη άδεια χρήσης, ώστε να μπορούν να τα αξιοποιήσουν οι αντίστοιχες υπηρεσίες.

Κομβικό ρόλο σε αυτή τη διαδικασία εξουσιοδότησης των χρηστών παίζει η στοιχειώδης υπηρεσία `AuthorizeUser`, η οποία αναλαμβάνει να ελέγξει κατά πόσο η πρόσβαση που αιτείται ο χρήστης πάνω στον πόρο επιτρέπεται από τον πάροχο ή το διανομέα του. Ο έλεγχος αυτός, σύμφωνα με το πρωτόκολλο της υπηρεσίας, γίνεται στη βάση των χαρακτηριστικών που διαθέτει και μπορούν να επιβεβαιωθούν με ω αντίστοιχων αδειών χρήσης ή ακόμα και σε ατομικό επίπεδο μέσω του ψηφιακού του πιστοποιητικού, ανάλογα με τους κανόνες που έχει ορίσει ο πάροχος του περιεχομένου.

Οι χρήστες αποκτούν χαρακτηριστικά που μπορούν να χρησιμοποιήσουν σε αυτό το πλαίσιο από διάφορους φορείς. Έτσι, μπορεί για παράδειγμα να προέρχονται από το κράτος όπου ανήκουν επιβεβαιώνοντας στοιχεία όπως υπηκοότητα και ηλικία, από τον οργανισμό όπου εργάζονται ή ακόμα και από τη λέσχη, στην οποία ανήκουν. Η ετερογένεια αυτών των φορέων έχει άμεση επίδραση στον τρόπο ανάθεσης των χαρακτηριστικών στους χρήστες. Για το λόγο αυτό, προτείνεται η χρήση της REL και της υπηρεσίας `CreateLicense` του MPEG-M για την έκδοση αδειών, οι οποίες περικλείουν και επιβεβαιώνουν την κατοχή ενός χαρακτηριστικού.

Συγκεκριμένα, κατά την έκδοση άδειας κατοχής ενός χαρακτηριστικού, η στοιχειώδης υπηρεσία `CreateLicense` πρώτα δημιουργεί το σχετικό κανόνα που περιγράφει ότι κάποιος χρήστης, ο οποίος αναγνωρίζεται μοναδικά είτε μέσω κάποιου αναγνωριστικού που διαθέτει μέσω της υπηρεσίας `IdentifyUser` είτε μέσω του δημόσιου κλειδιού του, όπως αυτό προκύπτει από το ψηφιακό πιστοποιητικό του. Στη συνέχεια, το αποτέλεσμα της παραπάνω ενέργειας υπογράφεται ψηφιακά, έτσι ώστε να μην μπορεί να αμφισβητηθεί η κατοχή του χαρακτηριστικού από το χρήστη.

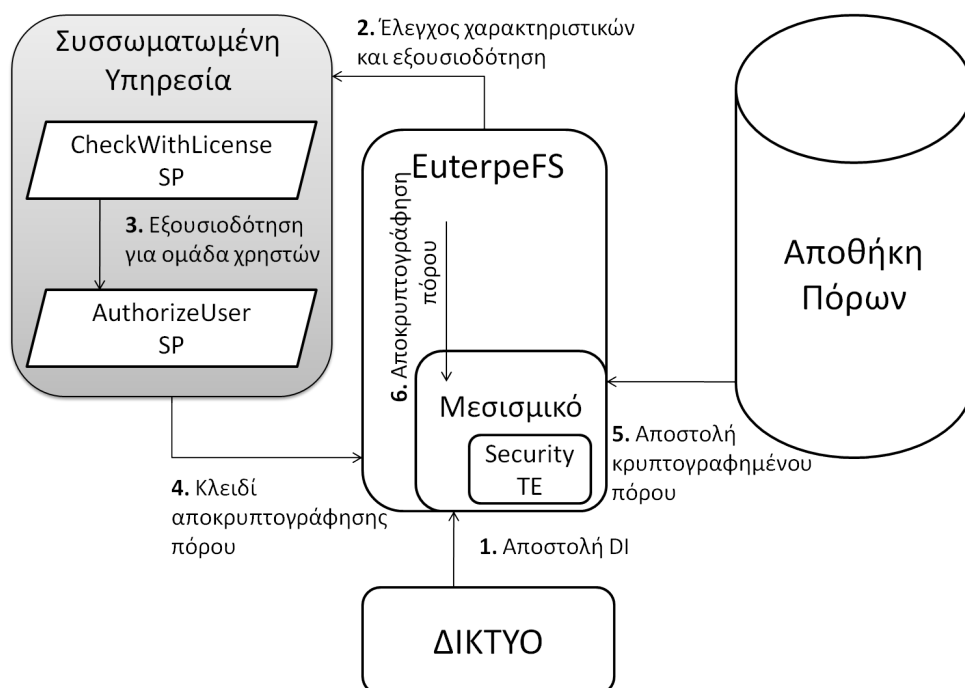
Η παραπάνω διαδικασία στην τρέχουσα έκδοση του προτύπου υποστηρίζεται με συμμετρική και ασύμμετρη κρυπτογραφία. Αν και θα μπορούσαν να χρησιμοποιηθούν πιο εξελιγμένα σχήματα, όπως αυτό που περιγράφεται στο [108], στα πλαίσια της παρούσας διατριβής θεωρήθηκε η τρέχουσα υλοποίηση που παρέχεται στο τρίτο μέρος του προτύπου [89].

Όταν πρόκειται να δημοσιεύσει κάποιος χρήστης ένα αρχείο στο σύστημα θεωρούμε ότι, στην τυπική περίπτωση, δημιουργεί και μια άδεια χρήσης. Αυτή η άδεια χρήσης ενσωματώνεται στο ψηφιακό αντικείμενο (είτε όπως είναι είτε ως αναφορά στην τοποθεσία όπου διατηρείται, όπως περιγράφηκε και στην προηγούμενη ενότητα), το οποίο αντιπροσωπεύει το φυσικό αρχείο στο EuterpeFS (υπενθυμίζεται ότι τα αρχεία του EuterpeFS είναι ψηφιακά αντικείμενα που δείχνουν στα φυσικά αρχεία). Στη συνέχεια, αυτή η άδεια χρήσης τίθεται σε ισχύ κρυπτογραφώντας το αρχείο. Σημειώνεται ότι το αρχείο θα είναι διαθέσιμο από αυτό το σημείο μόνο στην κρυπτογραφημένη του μορφή, για λόγους ασφάλειας.

λειας του συστήματος.

Δεδομένου ότι πρόκειται για κατανεμημένο σύστημα αρχείων, στη γενική περίπτωση ο φυσικός πόρος θα πρέπει να μπορεί να διαμοιράζεται μεταξύ πολλών χρηστών. Πλέον αυτού, λαμβάνοντας υπόψη ότι συνήθως υπάρχει ένα αντίγραφο του κρυπτογραφημένου αρχείου, στο οποίο θα δείχνει το εκάστοτε ψηφιακό αντικείμενο, γίνεται προφανές ότι το εν λόγω αρχείο θα πρέπει να κρυπτογραφείται συμμετρικά. Με αυτόν τον τρόπο, η εξουσιοδότηση συνίσταται στη διανομή του συμμετρικού κλειδιού στους κατάλληλους χρήστες. Το ρόλο αυτό στο MPEG-M τον έχει η στοιχειώδης υπηρεσία AuthorizeUser, όπως αναφέρθηκε και παραπάνω. Καθώς, όμως, η τελευταία δεν είναι σε θέση να γνωρίζει το κλειδί, αυτό περιλαμβάνεται στην άδεια χρήσης και κρυπτογραφείται με το δημόσιο κλειδί του παρόχου της υπηρεσίας AuthorizeUser που εμπιστεύεται ο πάροχος περιεχομένου.

Για να μπορέσει ο απομακρυσμένος χρήστης να αξιοποιήσει τα χαρακτηριστικά που διαθέτει για να πάρει πρόσβαση στην πληροφορία που επιθυμεί, πρέπει να βρίσκονται αποθηκευμένα στη μορφή των αδειών χρήσης με την οποία τα έλαβε. Για το λόγο αυτό, όταν λαμβάνει μια τέτοια άδεια, την αποθηκεύει χρησιμοποιώντας τη στοιχειώδη υπηρεσία StoreLicense. Δεδομένης της ευαισθησίας της εν λόγω πληροφορίας, θεωρείται ότι η υλοποίηση αυτής της υπηρεσίας είναι σε τοπικό επίπεδο και τα χαρακτηριστικά αποθηκεύονται είτε στο σκληρό δίσκο κρυπτογραφημένα είτε σε κάποια ασφαλή συσκευή, όπως είναι η έξυπνη κάρτα.



Σχήμα 45: Έλεγχος Πρόσβασης στο EuterpeFS

Όταν ένα απομακρυσμένο αρχείο του EuterpeFS (δηλαδή το ψηφιακό αντικείμενο που ενθυλακώνει τα μεταδεδομένα του φυσικού αρχείου) φτάσει στη συσκευή του τελικού

χρήστη, ακολουθείται η διαδικασία που περιγράφεται στο Σχήμα 45. Έτσι, αφού φτάσει το ψηφιακό αντικείμενο, ελέγχεται αν ο χρήστης έχει εξουσιοδότηση να πραγματοποιήσει την εκάστοτε ενέργεια. Η ενέργεια αυτή καθορίζεται από τη διαδικασία που θα καλέσει σε επίπεδο συστήματος αρχείων και μπορεί να είναι ανάγνωση (read) και εγγραφή (write). Σημειώνεται ότι μόνο αυτές οι δύο ενέργειες πραγματοποιούνται στον απομακρυσμένο πόρο, ενώ οι υπόλοιπες διαθέσιμες από το σύστημα αρχείων αφορούν το ψηφιακό αντικείμενο αυτό καθαυτό. Για να πραγματοποιηθεί ο έλεγχος, αποστέλλονται οι άδειες των χαρακτηριστικών που έχουν αποθηκευτεί, μαζί με την άδεια χρήσης που περιλαμβάνει το ληφθέν ψηφιακό αντικείμενο αρχείου στο πάροχο μιας ενοποιημένης υπηρεσίας που αναλαμβάνει να πραγματοποιήσει τον έλεγχο πρόσβασης. Το πρώτο βήμα αυτού του ελέγχου είναι να επιβεβαιωθεί ότι οι άδειες χρήσης των χαρακτηριστικών του χρήστη ισχύουν (διάρκεια χρήσης, αυθεντικοποίηση με ψηφιακή υπογραφή), κάτι το οποίο κάνει η στοιχειώδης υπηρεσία CheckWithLicense. Στη συνέχεια, και κατόπιν επιτυχούς πρώτου βήματος, η στοιχειώδης υπηρεσία AuthorizeUser θα ελέγξει κατά πόσο τα χαρακτηριστικά που δόθηκαν πληρούν τις προϋποθέσεις της πρόσβασης που αιτείται από το χρήστη και, σε περίπτωση επιτυχίας, επιστρέφει στο σύστημα αρχείων το συμμετρικό κλειδί. Σε αυτό το σημείο το EuterpeFS λαμβάνει τον πόρο που περιγράφει το ψηφιακό αντικείμενο (υπενθυμίζεται ότι μπορεί να πρόκειται ακόμα και για μια υπηρεσία ή μια ροή βίντεο και όχι απαραίτητα για ένα αρχείο) χρησιμοποιώντας το πρωτόκολλο που υποδεικνύεται στο αντίστοιχο στοιχείο. Ο πόρος αυτός καταφθάνει κρυπτογραφημένος, οπότε αποκρυπτογραφείται με το κλειδί που έλαβε νωρίτερα ο χρήστης από την ενοποιημένη υπηρεσία.

5.7 Οργάνωση Αρχείων στο EuterpeFS

Η βασική ιδέα που διαπνέει την οργάνωση αρχείων στο EuterpeFS πηγάζει από την παρατήρηση ότι οι χρήστες τείνουν να τοποθετούν τα αρχεία τους σε καταλόγους με βάση διάφορα σημασιολογικά κριτήρια, τα οποία φυσικά μπορεί να ποικίλουν κατά περίπτωση. Έτσι, σε ό,τι αφορά οργάνωση μουσικής για παράδειγμα, άλλοι χρήστες προτιμούν αρχειοθέτηση με βάση το είδος, άλλοι με βάση τον καλλιτέχνη, ενώ άλλοι ακόμα και αλφαβητικά με βάση το όνομα του τραγουδιού ή του καλλιτέχνη. Σε κάθε περίπτωση, όμως, τα αρχεία αυτά καθαυτά παραμένουν ίδια.

Ένα πρόβλημα που υπάρχει σε αυτές τις προσεγγίσεις είναι ότι ένα αρχείο μπορεί να ανήκει σημασιολογικά σε πολλές κατηγορίες. Έτσι, αν θεωρήσουμε έναν ερευνητή που θέλει να αρχειοθετήσει τις δημοσιεύσεις που συμβουλευεται για την έρευνά του, συχνά συναντά την περίπτωση ένα άρθρο να μπορεί να ταξινομηθεί σε πολλές περιοχές. Από την άλλη, η οργάνωση του περιεχομένου του σε ερευνητικές περιοχές μπορεί να διευκολύνει την αναζήτηση βιβλιογραφίας. Παρατηρείται, λοιπόν, εδώ μια σύγκρουση, η οποία συχνά επιλύεται με ύπαρξη αντιτύπων των αρχείων σε διαφορετικούς καταλόγους. Μάλιστα, αυτό πολλές φορές μπορεί να συμβεί τυχαία, δεδομένου ότι ο χρήστης μπορεί να

βρήκε ένα άρθρο για δεύτερη φορά ψάχνοντας για μια άλλη ερευνητική περιοχή, οπότε το αποθηκεύει και σε διαφορετικό κατάλογο.

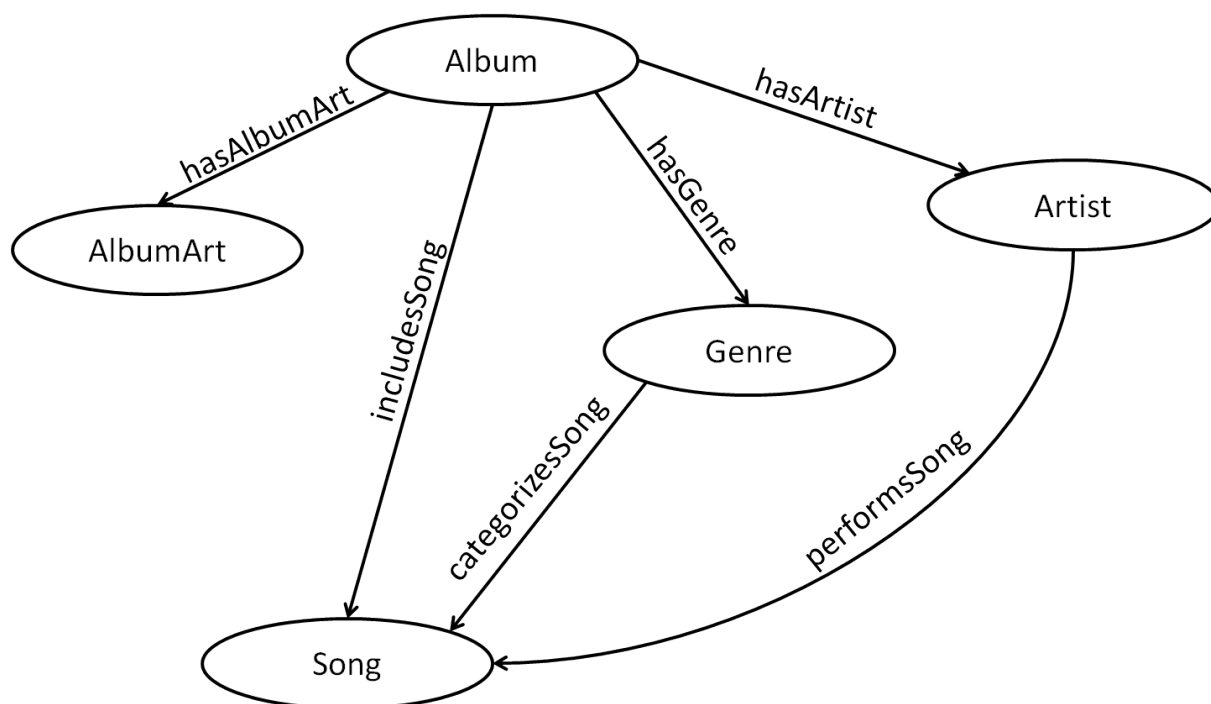
Ένας τρόπος που μπορεί να αντιμετωπιστεί το παραπάνω πρόβλημα είναι η αντιγραφή του αρχείου σε όλους τους καταλόγους που μπορεί να βρίσκεται. Ωστόσο, αυτό δημιουργεί προβλήματα συνέπειας, ενώ προκαλεί και κακή χρήση του διαθέσιμου χώρου στο σκληρό δίσκο, ειδικά στην περίπτωση πολυμεσικού περιεχομένου. Για να ξεπεραστούν αυτά τα ζητήματα, μια εναλλακτική λύση είναι η χρήση συμβολικών συνδέσμων, με το φυσικό αρχείο να βρίσκεται αποθηκευμένο σε ένα μοναδικό κατάλογο και όλοι οι υπόλοιποι να περιλαμβάνουν εικονικά αρχεία που δείχνουν στο πρωταρχικό. Η εμπειρία έχει δείξει, βέβαια, ότι κάτι τέτοιο σπάνια ακολουθείται στην πράξη.

Επιπλέον, όμως, των παραπάνω, ο χρήστης χρειάζεται να έχει και ευελιξία ως προς την οργάνωση των αρχείων. Έτσι, μπορεί ο τρόπος που έχει κάνει την αρχειοθέτηση να πρέπει να αλλάξει με την πάροδο του χρόνου, διότι άλλαξαν οι προτεραιότητές του, το αντικείμενο της εργασίας του κλπ. Για να γίνει κάτι τέτοιο, όμως, σε ένα παραδοσιακό σύστημα αρχείων απαιτείται χρόνος και κόπος από την πλευρά του χρήστη, ο οποίος θα πρέπει να φτιάξει από μόνος του τους καταλόγους που χρειάζεται και στη συνέχεια, να μεταφέρει τα αρχεία του στη νέα δομή.

Αυτά τα ζητήματα γίνονται ακόμα μεγαλύτερα σε ένα κατανεμημένο περιβάλλον, όπου ο αριθμός των αρχείων που πρέπει να διαχειριστεί ο χρήστης είναι μεγαλύτερος, ενώ επιπλέον θα πρέπει συνήθως να βασιστεί και στη δομή που έχουν ήδη επιλέξει οι πάροχοι του περιεχομένου. Λαμβάνοντας υπόψη τα παραπάνω και σε συνδυασμό με την πρόοδο που έχει επιτελεστεί στον τομέα των σημασιολογικών (βλ. σημασιολογικός ιστός), το EuterpeFS προτείνει ένα δυναμικό τρόπο οργάνωσης των αρχείων στο τοπικό σύστημα αρχείων, με βάση οντολογίες που παρέχει ο χρήστης, οι οποίες υπαγορεύουν τη θέση των αρχείων με βάση τα μεταδεδομένα τους.

Συγκεκριμένα κάθε οντολογική έννοια (κλάση) αναπαρίσταται με έναν κατάλογο στο σύστημα αρχείων, οπότε η διάτρεξη των καταλόγων αντιστοιχεί στη διάτρεξη των αντίστοιχων οντολογικών εννοιών. Τα αρχεία τοποθετούνται στους καταλόγους με βάση τη (σημασιολογική) περιγραφή τους. Έτσι, κάθε ψηφιακό αντικείμενο αρχείου που εμφανίζεται σε κάποιον κατάλογο θα πρέπει να περιλαμβάνει μεταδεδομένα, τα οποία απαντούν στο ερώτημα που προκύπτει από την αντίστοιχη οντολογική κλάση.

Η αναπαράσταση των καταλόγων δε γίνεται σε επίπεδο εννοιών, αλλά στιγμιότυπων τους, όπως αυτά προκύπτουν από την εκάστοτε σημασιολογική περιγραφή, καθώς και ιδιοτήτων, οι οποίες συνδέουν τις έννοιες. Έτσι, εφόσον κάποιο στιγμιότυπο αναφέρεται απλά ως όνομα και όχι ως αναγνωριστικό ενός ψηφιακού αντικειμένου, αυτό θα αποτελεί κατάλογο, ενώ στην αντίθετη περίπτωση ύπαρξης αναγνωριστικού, αυτό θα είναι αρχείο, το οποίο τοποθετείται στον κατάλογο που αντιστοιχεί στον τύπο του. Τελικά, στην πράξη, το αν μια οντότητα στο σύστημα αρχείων αναπαρίσταται ως αρχείο ή κατάλογος, εξαρτάται με το αν υπάρχει ψηφιακό αντικείμενο που έχει τύπο οντολογική έννοια. Το Σχήμα 46



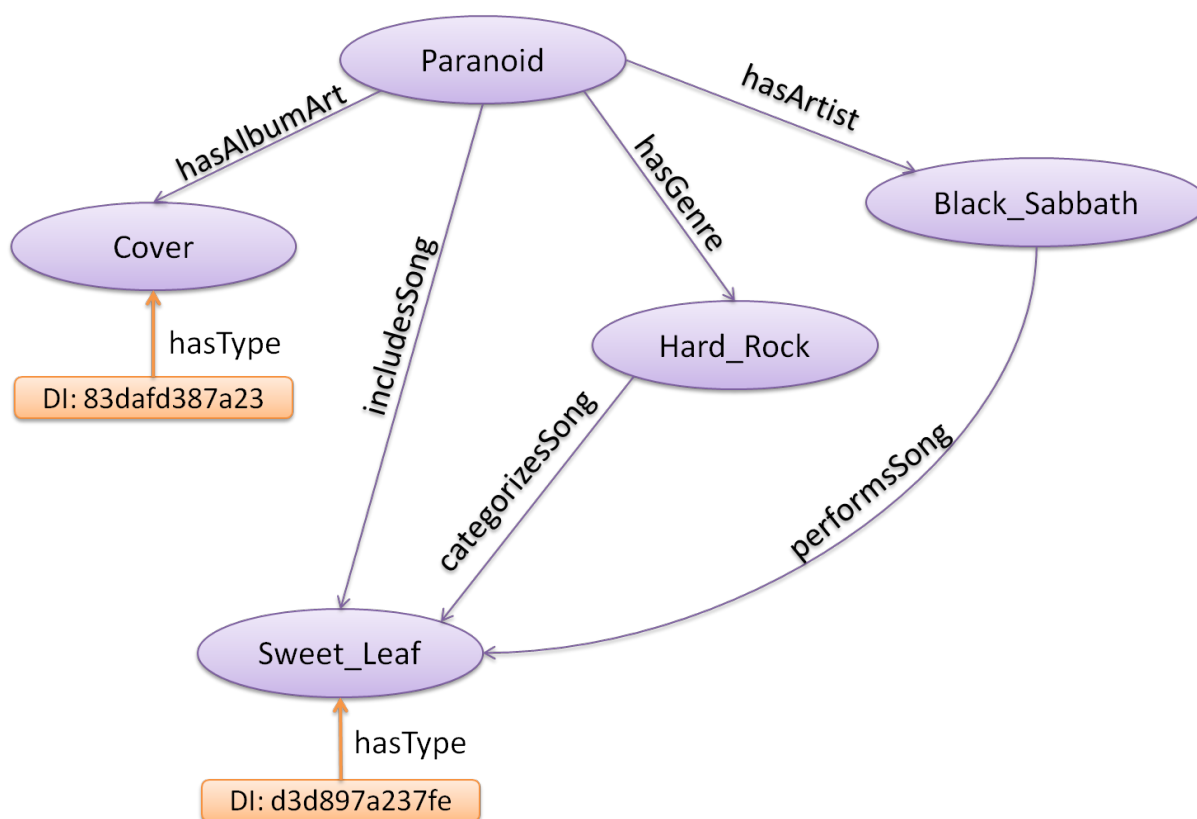
Σχήμα 46: Παράδειγμα Μουσικής Οντολογίας - Επίπεδο Εννοιών

παρουσιάζει ένα παράδειγμα μουσικής οντολογίας σε εννοιολογικό επίπεδο που μπορεί να χρησιμοποιηθεί για την οργάνωση των καταλόγων.

Περνώντας σε επίπεδο στιγμιότυπων, το Σχήμα 47 παρουσιάζει την αντίστοιχη οντολογία. Όπως φαίνεται υπάρχουν δύο στιγμιότυπα, ένα για την έννοια AlbumArt (εξώφυλλο) και ένα για την έννοια Song (τραγουδι Sweet Leaf), στα οποία έχουν αντιστοιχηθεί μέσω της σημασιολογικής τους περιγραφής δύο ψηφιακά αντικείμενα αρχείου. Μετά την προσάρτηση στο τοπικό σύστημα αρχείων, το αποτέλεσμα θα είναι να υπάρχουν καταλόγοι για όλα τα στιγμιότυπα που δεν έχουν αντιστοιχηθεί σε ψηφιακά αντικείμενα, καθώς και όλες τις ιδιότητες που τα συνδέουν. Για τα στιγμιότυπα των δύο εννοιών που προαναφέρθηκαν θα υπάρχουν μόνο τα ψηφιακά αντικείμενα αρχείου που βρέθηκαν διαθέσιμα. Έτσι, για να φτάσει ο χρήστης στο τραγουδι Sweet Leaf, μπορεί να φτάσει από τις εξής εναλλακτικές διαδρομές:

- Paranoid/songs/Sweet Leaf
- Paranoid/artists/Black Sabbath/songs/Sweet Leaf
- Paranoid/genres/Hard Rock/songs/Sweet Leaf

Προκειμένου να μην επαναλαμβάνονται τα αρχεία σε κάθε κατάλογο και με δεδομένο ότι το παραπάνω πρόκειται για απλή απεικόνιση, το EuterpeFS διατηρεί ευρετήριο με τα ψηφιακά αντικείμενα αρχείου, όπου έχει κάνει συνδρομή. Αυτό το ευρετήριο βρίσκεται αποθηκευμένο στο τοπικό σύστημα αρχείων που χρησιμοποιεί έτσι κι αλλιώς και κάθε



Σχήμα 47: Παράδειγμα Μουσικής Οντολογίας - Επίπεδο Στιγμιότυπων

φορά που επιχειρεί να ανοίξει έναν κατάλογο, πραγματοποιείται το αντίστοιχο ερώτημα και παίρνει τα ψηφιακά αντικείμενα που διαθέτουν πληροφορία για την έννοια που αναπαριστά ο κατάλογος. Το ίδιο συμβαίνει και όταν επιχειρεί να ανοίξει κάποιο αρχείο, οπότε επιστρέφεται το αποθηκευμένο ψηφιακό αντικείμενο και οι πράξεις διενεργούνται πάνω σε αυτό.

Όταν ο χρήστης δημιουργεί ένα νέο κατάλογο, η ενέργεια αυτή ισοδυναμεί με τη δημιουργία στιγμιότυπου της οντολογικής έννοιας στην οποία βρίσκεται. Έτσι, στο προηγούμενο παράδειγμα, έστω ότι δημιουργείται ένας νέος κατάλογος Heavy Metal μέσα στον κατάλογο genres. Αυτό θα προκαλέσει ενημέρωση της οντολογίας ότι υπάρχει κι ένα νέο στιγμιότυπο της έννοιας Genre διαθέσιμο, με όνομα Heavy Metal.

5.8 Δημοσίευση Αρχείων στο EuterpeFS

Η δημοσίευση ενός αρχείου στο EuterpeFS σημαίνει και ταυτόχρονα διαμοιρασμό του με τους υπόλοιπους χρήστες. Η διαδικασία που ακολουθείται ισοδυναμεί με αυτήν που περιγράφηκε ήδη με τη μορφή ενοποιημένης υπηρεσίας στην Ενότητα 4.2.4. Έτσι, πρώτα δημιουργείται ένα ψηφιακό αντικείμενο αρχείου, το οποίο περιλαμβάνει σημασιολογική περιγραφή, όπως αυτή προκύπτει από τον τρέχοντα κατάλογο όπου βρίσκεται το προς δη-

μοσίευση αρχείο. Στη συνέχεια, προστίθεται ένα μοναδικό αναγνωριστικό και η άδεια χρήσης, εφόσον το ζητήσει ρητά ο χρήστης.

Στην περίπτωση που ο χρήστης επιθυμεί να δημοσιεύσει ένα τοπικό αρχείο, αυτό που έχει να κάνει αρχικά είναι να βρει τη διαδρομή καταλόγων που ανταποκρίνεται στη σημασιολογική του περιγραφή. Αν δεν υπάρχουν οι κατάλογοι, τότε τους δημιουργεί, με ταυτόχρονη δημιουργία των αντίστοιχων στιγμιοτύπων, όπως περιγράφηκε νωρίτερα στην Ενότητα 5.7. Στη συνέχεια αντιγράφει το αρχείο προς δημοσίευση στον κατάλογο αυτό. Για την περίπτωση των τοπικών αρχείων, ο χρήστης θα πρέπει να διαθέτει κάποια υπηρεσία που να υποστηρίζει την αποστολή του. Στην τυπική περίπτωση, αυτό θα μπορούσε να είναι μια υπηρεσία HTTP. Ωστόσο, λαμβάνοντας υπόψη και όσα παρουσιάστηκαν στο Κεφάλαιο 2, θα μπορούσε να είναι και μια υπηρεσία διαμοιρασμού περιεχομένου χρησιμοποιώντας το πρωτόκολλο BitTorrent ή αν πρόκειται για βίντεο, υπηρεσία ροής του.

Για να γίνει η αντιγραφή, πραγματοποιείται κλήση συστήματος στην εντολή `write`, η οποία υλοποιείται από το EuterpeFS με τον εξής τρόπο: Αρχικά το φυσικό αρχείο γίνεται διαθέσιμο για μεταφορά, κάνοντας τις ενέργειες που υπαγορεύει το πρωτόκολλο της επιλογής. Για παράδειγμα, εφόσον πρόκειται να χρησιμοποιηθεί το HTTP, γίνεται η κατάλληλη διαμόρφωση του εξυπηρετητή ιστού, ώστε να μπορεί να διαθέσει το εν λόγω αρχείο. Στη συνέχεια, δημιουργείται η σημασιολογική περιγραφή που αντιστοιχεί στον κατάλογο όπου τοποθετήθηκε το αρχείο. Κατόπιν, δημιουργείται το ψηφιακό αντικείμενο αρχείου, το οποίο περιλαμβάνει την παραπάνω σημασιολογική περιγραφή, μια αναφορά στο URL με το οποίο υποστηρίζεται η αποστολή του αρχείου, καθώς και πληροφορίες που αφορούν τη δομή `stat`, όπως το μέγεθος του φυσικού αρχείου.

Μετά το πέρας αυτής της διαδικασίας, παρουσιάζεται ως αρχείο στον τοπικό κατάλογο του EuterpeFS το ψηφιακό αντικείμενο του προς δημοσίευση αρχείου, το οποίο μπαίνει σε μια ουρά αναμένοντας να λάβει αναγνωριστικό και να δημοσιευτεί. Η χρόνος αναμονής είναι παράμετρος του συστήματος και ο ρόλος της είναι να δώσει ένα χρονικό διάστημα στο χρήστη να πραγματοποιήσει περαιτέρω ενέργειες στο ψηφιακό αντικείμενο πριν τη δημοσίευση, όπως η προσθήκη άδειας χρήσης ή άλλων μεταδεδομένων.

Ειδικά για την τελευταία περίπτωση, είναι πιθανό ο χρήστης να επιθυμεί να εισάγει και επιπλέον μεταδεδομένα καθώς θεωρεί ότι υπάρχουν επιπλέον κατάλογοι όπου θα μπορούσε να βρίσκεται. Έτσι, ανατρέχοντας στο προηγούμενο παράδειγμα του τραγουδιού Sweet Leaf, ο δημοσιεύων χρήστης θα έπρεπε να αντιγράψει το πραγματικό αρχείο στους τρεις καταλόγους που αναφέρθηκαν. Η αντιγραφή ψηφιακού αντικειμένου στο EuterpeFS στην πραγματικότητα δημιουργεί ένα συμβολικό σύνδεσμο στο αρχικό ψηφιακό αντικείμενο, το οποίο όμως ανανεώνεται με επιπλέον περιγραφικά μεταδεδομένα, όπως αυτά προκύπτουν από τον εκάστοτε κατάλογο.

Για τη δημιουργία της άδειας χρήσης, χρησιμοποιούνται οι εντολές `chown` και `chmod` που παρέχονται έτσι κι αλλιώς από το λειτουργικό σύστημα. Δεδομένου, όμως, ότι οι παραδοσιακές τους υλοποιήσεις αφορούν μεμονωμένους χρήστες και ομάδες τους, όπως αυ-

τές δίνονται από το τοπικό σύστημα, είναι απαραίτητο να υπάρχει ένα επιπλέον αρχείο ιδιοτήτων, το οποίο περιλαμβάνει τη λίστα με τους χρήστες και τις ομάδες τους στο κατανεμημένο περιβάλλον, όπου πρόκειται να λειτουργήσει το EuterpeFS. Πλέον αυτών, είναι επίσης απαραίτητη η ύπαρξη ενός αρχείου ρημάτων, κατά τα πρότυπα του λεξικού δικαιωμάτων πάνω στα δεδομένα του MPEG-21 (RDD) [71], ώστε να μπορεί το σύστημα αρχείων να ανατρέξει εκεί για τον τρόπο πρόσβασης που υποδεικνύει ο χρήστης. Τα δικαιώματα αυτά είναι αριθμημένα και, όπως θα εξηγηθεί παρακάτω, δεν μπορούν να είναι περισσότερα από 512 διαφορετικά ρήματα.

Με την εντολή `chown`, η άδεια χρήσης ενημερώνεται ως προς το πεδίο *principal*, το οποίο υποδεικνύει το χρήστη ή την ομάδα χρηστών όπου απευθύνονται τα προνόμια που πρόκειται να επακολουθήσουν. Στη συνέχεια, με τη χρήση της εντολής `chmod` δίνεται ρητά το προνόμιο, το οποίο έχει ο εκάστοτε χρήστης (ή ομάδα χρηστών) στο αρχείο. Ο οκταδικός αριθμός που δίνεται σε αυτήν περίπτωση αποτελεί στην πραγματικότητα τον αριθμό ευρετηρίου που έχει λάβει στο αντίστοιχο αρχείο ιδιοτήτων με τα πιθανά προνόμια που μπορούν να ανατεθούν. Αυτός είναι και ο λόγος για τον οποίο ο αριθμός των διαθέσιμων προνομίων δεν μπορεί να ξεπεράσει το 512 ($8^3 = 512$).

Όταν έχουν πια κληθεί και οι δύο εντολές (`chown` και `chmod`) και η άδεια χρήσης έχει δημιουργηθεί, γίνεται η κρυπτογράφηση του αρχείου στο χώρο όπου τοποθετήθηκε για να γίνει διαθέσιμο για μεταφορά, ώστε να είναι πια προστατευμένο. Αν πρόκειται για άδεια χρήσης που αφορά μεμονωμένο χρήστη, τότε χρησιμοποιείται το δημόσιο κλειδί του, που βρίσκεται στο αρχείο ιδιοτήτων. Αν, από την άλλη, αφορά ομάδα χρηστών, τότε παράγεται συμμετρικό κλειδί, το οποίο και τοποθετείται κρυπτογραφημένο στην άδεια χρήσης, ακολουθώντας τη διαδικασία που περιγράφηκε νωρίτερα στην Ενότητα 5.6.

Εφόσον το προς δημοσίευση αρχείο είναι κάποιο απομακρυσμένο, στο οποίο έχει πρόσβαση ο χρήστης μέσω κάποιου άλλου συστήματος αρχείων (π.χ. NFS), τότε οι δυνατότητες που διαθέτει είναι περιορισμένες, εφόσον επιθυμεί να συνεχίσει να το διαθέτει μέσω του άλλου συστήματος αρχείων και να μην το αντιγράψει στον τοπικό του δίσκο. Εκεί, το μόνο που μπορεί να κάνει είναι να προσθέσει σημασιολογικές περιγραφές και όχι άδειες χρήσης, καθώς δεν μπορεί να κρυπτογραφήσει τον απομακρυσμένο πόρο. Η αναφορά στο φυσικό πόρο που υπάρχει στο ψηφιακό αντικείμενο είναι στην πραγματικότητα το URL που προκύπτει από τον τρόπο που διαμοιράζεται σε πρωταρχικό στάδιο.

Όταν παρέλθει το χρονικό διάστημα αναμονής για δημοσίευση και, εφόσον δεν είναι ενεργή κάποια από τις παραπάνω διαδικασίες, καλείται η στοιχειώδης υπηρεσία του MPEG-M `IdentifyContent`, προκειμένου το ψηφιακό αντικείμενο να λάβει ένα μοναδικό αναγνωριστικό. Στη συνέχεια, καλείται η `PostContent`, προκειμένου αυτό το ψηφιακό αντικείμενο να δημοσιευτεί και να γίνει διαθέσιμο στους υπόλοιπους χρήστες και το ψηφιακό αντικείμενο αρχείου σημειώνεται ως δημοσιευμένο.

Σε περίπτωση που το ψηφιακό αντικείμενο αρχείου διαγραφεί από κάποια τοποθεσία, τότε η περιγραφή του θα πρέπει να αλλάξει, ενώ αν διαγραφεί από όλες, θα πρέπει να

διαγραφεί εντελώς και να πάψει να είναι δημοσιευμένο. Αντίστοιχα και όταν γίνει κάποια αλλαγή στην άδεια χρήσης, θα πρέπει το δημοσιευμένο ψηφιακό αντικείμενο να ενημερωθεί. Για το σκοπό αυτό, όταν γίνεται κάποια τέτοια αλλαγή, το ψηφιακό αντικείμενο σημειώνεται ότι βρίσκεται υπό επεξεργασία και όταν παρέλθει το χρονικό διάστημα δημοσίευσης, καλείται αρχικά η στοιχειώδης υπηρεσία `RevokeContent` για να ανακληθεί η προηγούμενη δημοσίευση και στη συνέχεια ακολουθείται ξανά η προαναφερθείσα διαδικασία δημοσίευσης με τις στοιχειώδεις υπηρεσίες `IdentifyContent` και `PostContent` να καλούνται και πάλι για το νέο ψηφιακό αντικείμενο.

Ωστόσο, φροντίδα θα πρέπει να λαμβάνεται και όταν υπάρχει αλλαγή στο φυσικό αρχείο. Έτσι, το EuterpeFS επηρεάζεται από τυχών αλλαγή του μεγέθους του αρχείου, του τρόπου μεταφοράς και, φυσικά, από τη διαγραφή του. Δεδομένου, όμως, ότι αυτές οι αλλαγές πραγματοποιούνται σε διαφορετικό σύστημα αρχείων, το EuterpeFS δεν μπορεί παρά να ελέγχει περιοδικά τα δημοσιευμένα φυσικά αρχεία για τέτοιες αλλαγές, ώστε να προχωρήσει σε επαναδημοσίευση του αρχείου (όπως περιγράφηκε προηγουμένως) ή σε οριστική απόσυρση της δημοσίευσης (στην περίπτωση διαγραφής του φυσικού αρχείου).

5.9 Συνδρομή σε Αρχεία στο EuterpeFS

Η συνδρομή σε αρχεία γίνεται όταν ο χρήστης επιχειρεί προσάρτηση του κατανεμημένου συστήματος αρχείων EuterpeFS σε κάποιον τοπικό κατάλογο. Τα απομακρυσμένα αρχεία εντοπίζονται με βάση τα ενδιαφέροντα του εκάστοτε χρήστη, όπως αυτά εκφράζονται μέσω της δημιουργίας ενός ψηφιακού αντικειμένου συνδρομής (βλ. Ενότητα 4.2.4) με τη σημασιολογική περιγραφή αυτών των ενδιαφερόντων. Αυτό το ψηφιακό αντικείμενο δίνεται ως όρισμα στην εντολή προσάρτησης `mount` και το χρησιμοποιεί το EuterpeFS για να το τροφοδοτήσει στη στοιχειώδη υπηρεσία `PostContent`, ώστε να πραγματοποιηθεί η συνδρομή βασιζόμενη στο περιεχόμενο. Επιπρόσθετα, δίνεται ως όρισμα το αρχείο της οντολογίας που υποδεικνύει την ιεραρχία των καταλόγων, με βάση τη σημασιολογική περιγραφή των διαμοιραζόμενων αρχείων.

Όπως είναι προφανές από αυτή τη διαδικασία, το EuterpeFS, σε αντίθεση με τα παραδοσιακά κατανεμημένα συστήματα αρχείων που προσαρτούν ολόκληρο έναν απομακρυσμένο κατάλογο, δημιουργεί αφενός τους καταλόγους που ορίζονται στην οντολογία που δίνεται ως όρισμα, ενώ παρέχει διαφάνεια τοποθεσίας μόνο σε εκείνα τα αρχεία που ενδιαφέρουν το χρήστη. Με αυτόν τον τρόπο ακολουθείται μια προσέγγιση που επικεντρώνεται στην προσωποποιημένη οργάνωση του περιεχομένου που ενδιαφέρει το χρήστη, και μόνο εκείνου.

Κάθε δημοσιευμένο αρχείο, του οποίου η σημασιολογική περιγραφή απαντά στο ερώτημα του ψηφιακού αντικειμένου συνδρομής μεταφέρεται στη συσκευή του χρήστη και αποθηκεύεται σε ειδική τοποθεσία. Όταν ο χρήστης προσπελαύνει το προσαρτημένο

σύστημα αρχείων και για κάθε κατάλογο όπου μπαίνει, πραγματοποιείται το αντίστοιχο ερώτημα στα ληφθέντα ψηφιακά αντικείμενα. Για λόγους ενίσχυσης της αποδοτικότητας και αποφυγής ερωτημάτων, διατηρείται επιπλέον δομή ευρετηρίου, η οποία περιλαμβάνει αναφορές σε ήδη υπάρχοντα ψηφιακά αντικείμενα ανά κατάλογο, ώστε κάθε φορά που γίνεται προσπέλαση του καταλόγου, να πραγματοποιείται το σχετικό ερώτημα μόνο στα νέα ψηφιακά αντικείμενα.

Για την αναπαράσταση του αρχείου σε κάποιον κατάλογο χρησιμοποιείται η πληροφορία που ενθυλακώνεται στο ψηφιακό αντικείμενο για τη δομή stat. Όπως αναφέρθηκε και σε προηγούμενη ενότητα, είναι απαραίτητο εκεί να δίνεται το μέγεθος του πραγματικού αρχείου, ώστε η εντολή συστήματος read να είναι σε θέση να γνωρίζει πώς θα το προσπελάσει. Επιπλέον του μεγέθους, μπορεί να υπάρχει το όνομα αρχείου, με εναλλακτική επιλογή τη χρήση του αναγνωριστικού του ψηφιακού αντικειμένου ως ονόματος. Το τελευταίο δεν προτείνεται, δεδομένου ότι το αναγνωριστικό είναι τις περισσότερες φορές μια συμβολοσειρά άνευ νοήματος.

Όταν ο χρήστης επιχειρήσει να ανοίξει ένα ψηφιακό αντικείμενο αρχείου, τότε ελέγχεται πρώτα το είδος του αρχείου που πρόκειται να ανοίξει. Ανάλογα με το είδος του αρχείου μπορεί να διαφέρει και η ενέργεια σε επίπεδο ελέγχου πρόσβασης (π.χ. αν το φυσικό αρχείο είναι βίντεο, τότε το ρήμα είναι "play"). Στη συνέχεια, επιλέγεται το πρόγραμμα που θα χρησιμοποιηθεί για να προσπελάσει το αρχείο (π.χ. εφαρμογή αναπαραγωγής βίντεο). Αν αυτή η εφαρμογή είναι συμβατή με MPEG-M, τότε μπορεί να τροφοδοτήσει και το σχετικό ρήμα. Εναλλακτικά (που αποτελεί και τη γενική περίπτωση), χρησιμοποιείται το "read", αφού πρόκειται στην πραγματικότητα για ανάγνωση.

Λαμβάνοντας υπόψη ότι το φυσικό αρχείο θα είναι προστατευμένο, ήτοι κρυπτογραφημένο, ο χρήστης θα πρέπει να αιτηθεί το κλειδί της αποκρυπτογράφησης. Για το λόγο αυτό, θα ελέγξει την άδεια και, εφόσον αυτή αναφέρεται ρητά σε εκείνον, τότε μπορεί να χρησιμοποιήσει το ιδιωτικό του κλειδί, καθώς το φυσικό αρχείο θα έχει κρυπτογραφηθεί με το αντίστοιχο δημόσιο, όπως περιγράφηκε στην προηγούμενη ενότητα. Σε περίπτωση που πρόκειται να γίνει χρήση των χαρακτηριστικών που διαθέτει ο χρήστης προκειμένου να εξασφαλίσει την πρόσβαση, ακολουθείται η διαδικασία που αναλύθηκε στην Ενότητα 5.6. Στην περίπτωση επιτυχούς εξουσιοδότησης, ο χρήστης θα λάβει το συμμετρικό κλειδί και θα μπορέσει να αποκρυπτογραφήσει το φυσικό αρχείο. Σε αυτό το σημείο, ξεκινά και η μεταφορά του τελευταίου, με βάση το πρωτόκολλο που έχει χρησιμοποιηθεί.

Κεφάλαιο 6

Συμπεράσματα και Μελλοντικό Έργο

6.1 Σύνοψη και Συμπεράσματα

Η διαθέσιμη πληροφορία στο διαδίκτυο αυξάνεται συνεχώς. Ήδη, το 2008 η Google διαπίστωσε ότι ο αριθμός των ιστοσελίδων στον παγκόσμιο ιστό έφτανε το ένα τρισεκατομμύριο [109], ενώ η αύξησή του ακολουθεί εκθετικό ρυθμό [110]. Επιπλέον, η κίνηση που κυριαρχεί, αλλά και αναμένεται να συνεχίσει να κυριαρχεί τα επόμενα χρόνια, στο διαδίκτυο αφορά ομότιμα πρωτόκολλα διαμοιρασμού αρχείων, με επικρατέστερο το BitTorrent, και ροή βίντεο [24] [28].

Λαμβάνοντας υπόψη αυτά τα δεδομένα, σε συνδυασμό με την ετερογένεια των συσκευών και των πλατφορμών των χρηστών, η παρούσα διατριβή επιχειρεί να αποτελέσει μια ολοκληρωμένη λύση που απευθύνεται σε όλο τον κύκλο ζωής του περιεχομένου στο δίκτυο: από τη δημοσίευση, μέχρι την ανακάλυψη και την παράδοσή του στον τελικό χρήστη. Ειδικά σε ό,τι αφορά την παράδοσή του, δόθηκε έμφαση στα ομότιμα πρωτόκολλα και στη ροή βίντεο που αποτελούν και τη μεγαλύτερη πρόκληση στο προσεχές μέλλον. Έτσι, οι βασικοί άξονες της διατριβής συνίστανται στα ακόλουθα σημεία:

1. Δημοσίευση περιεχομένου για παράδοση από πληροφοριοκεντρικό δίκτυο. Τα πληροφοριοκεντρικά δίκτυα αποτελούν μια νέα επαναστατική προσέγγιση στον τρόπο, με τον οποίο δρομολογείται το περιεχόμενο στο επίπεδο του δικτύου. Σε αντίθεση με τα σύγχρονα δίκτυα, όπου πρέπει πρώτα να εξευρεθεί η διεύθυνση του κόμβου εξυπηρέτησης του περιεχομένου και, στη συνέχεια, να ζητηθεί το περιεχόμενο από αυτόν, με βάση κάποιο πρωτόκολλο όπως το HTTP, τα πληροφοριοκεντρικά δίκτυα δίνουν τη δυνατότητα δρομολόγησης αιτήσεων με βάση το όνομα του περιεχομένου, χωρίς ο αιτών να γνωρίζει από πού πρέπει να ζητηθεί, αφού αυτό αποτελεί έργο του δικτύου. Αυτός ο τρόπος δρομολόγησης με βάση το όνομα επιτρέπει στο δίκτυο να

- γνωρίζει ακριβώς την πληροφορία που ζητείται, οδηγώντας σε αποτελεσματικότερη χρήση του εύρους ζώνης αφενός με έλεγχο των αιτήσεων που προωθούνται προς τον εκάστοτε κόμβο εξυπηρέτησης περιεχομένου και αφετέρου με τη χρήση κρυφής μνήμης, ώστε κάποιες αιτήσεις να εξυπηρετούνται από το ίδιο το δίκτυο. Αυτές τις ιδιότητες αξιοποιούν δύο προτεινόμενα πρωτόκολλα για ομότιμα δίκτυα και ροή βίντεο, επεκτείνοντας τους υπάρχοντες μηχανισμούς των πληροφοριοκεντρικών δικτύων.
2. Παράδοση περιεχομένου σε πληροφοριοκεντρικό δίκτυο, με έμφαση στη ροή βίντεο και στα ομότιμα δίκτυα διαμοιρασμού αρχείων. Αυτές οι δύο κατηγορίες αποτέλεσαν βασικό πεδίο έρευνας της παρούσας διατριβής, στα πλαίσια των πληροφοριοκεντρικών δικτύων. Έτσι, παρουσιάζονται εγγενείς σχεδιαστικές αδυναμίες των υπαρχουσών αρχιτεκτονικών, καθώς και λύσεις, οι οποίες αυξάνουν την απόδοση του συστήματος. Η αύξηση αυτή αποτυπώνεται με πειραματικά αποτελέσματα, τα οποία προέκυψαν από προσομοιώσεις που πραγματοποιήθηκαν με χρήση ενός προσομοιωτή που αναπτύχθηκε ειδικά για πληροφοριοκεντρικά δίκτυα και ο οποίος επίσης παρουσιάζεται στο αντίστοιχο κεφάλαιο.
 3. Ανακάλυψη του περιεχομένου, με χρήση πρωτοκόλλου δημοσίευσης/συνδρομής βασισμένο στο περιεχόμενο. Ακολουθώντας το μοντέλο των επιδημικών αλγορίθμων, οι οποίοι χρησιμοποιούνται κατά κόρον στη βιβλιογραφία ως βάση για διάφορα πρωτόκολλα που έχουν προταθεί, το προτεινόμενο πρωτόκολλο CCPS βελτιώνει την απόδοση του συστήματος τόσο ως προς την ικανότητα κλιμάκωσης όσο και ως προς την εξισορρόπηση φόρτου, σε σχέση με ένα τυπικό επιδημικό πρωτόκολλο. Τα αποτελέσματα που δίνονται βασίζονται αρχικά σε θεωρητική ανάλυση του συστήματος και στη συνέχεια επιβεβαιώνονται από προσομοιώσεις που πραγματοποιήθηκαν με βάση υλοποιήσεις των εμπλεκόμενων πρωτοκόλλων στον προσομοιωτή PeerSim.
 4. Προσωποποιημένη οργάνωση κατανεμημένου περιεχομένου σε τοπικό επίπεδο. Η πιο συνηθισμένη μορφή απεικόνισης της πληροφορίας σε τοπικό επίπεδο είναι το αρχείο. Ωστόσο, σε ένα κατανεμημένο περιβάλλον, όπου οι χρήστες διαμοιράζονται περιεχόμενο σε μορφή αρχείων, η επιλογή της πληροφορίας που ενδιαφέρει τον κάθε χρήστη καθώς και η οργάνωσή της στο τοπικό σύστημα αρχείων αποτελεί πρόκληση. Με το προτεινόμενο κατανεμημένο σύστημα αρχείων EuterpeFS δίνεται η δυνατότητα στους χρήστες αφενός να γίνονται συνδρομητές σε περιεχόμενο που τους ενδιαφέρει μέσω συνδρομών βασισμένων στο περιεχόμενο, χρησιμοποιώντας για παράδειγμα το πρωτόκολλο δημοσίευσης/συνδρομής που επίσης παρουσιάζεται, και αφετέρου να οργανώνουν τα αρχεία σε τοπικούς καταλόγους με βάση τη σημασιολογική περιγραφή τους. Επιπλέον, η προστασία της πληροφορίας γίνεται με χρήση πρότυπων τεχνολογιών, προκειμένου να επιτυγχάνεται έλεγχος πρόσβασης σε κατανεμημένο περιβάλλον με ομοιογενή τρόπο.
 5. Υποστήριξη ετερογένειας με χρήση πρότυπων τεχνολογιών. Οι τεχνολογίες αυτές, οι οποίες συμπεριλαμβάνονται στο μεσισμικό του MPEG (MPEG-M) αφορούν σε έναν

ενιαίο τρόπο ενθουλάκωσης του περιεχομένου, καθώς και πληροφορίας που το αφορά και έχει να κάνει με τον έλεγχο πρόσβασης σε αυτό, την περιγραφή του, τον τρόπο πρόσβασης σε αυτό κλπ. Στις τεχνολογίες αυτού του προτύπου έχει βασιστεί το κατανεμημένο σύστημα αρχείων EuterpeFS, ενώ και το πρωτόκολλο CCPS μπορεί να υλοποιηθεί χρησιμοποιώντας τις αντίστοιχες λειτουργίες που έτσι κι αλλιώς ορίζει το πρότυπο. Με αυτόν τον τρόπο, το MPEG-M αποτελεί το συνδεδετικό κρίκο που ενώνει και τα τρία επίπεδα της διατριβής: παράδοση, ανακάλυψη και οργάνωση του περιεχομένου.

6.2 Μελλοντικές Ερευνητικές Κατευθύνσεις

Το τρέχον και μελλοντικό έργο της παρούσας διατριβής αφορά κυρίως τον τομέα της ανακάλυψης και της παράδοσης του περιεχομένου. Ωστόσο, ακόμα και τα κομμάτια που αφορούν την οργάνωση του περιεχομένου καθώς και την ενοποίηση του συστήματος επιδέχονται βελτιώσεων, ώστε να ληφθούν υπόψη επιπλέον παράμετροι που αφορούν το πλαίσιο που πρόκειται να διαμορφωθεί με την εισαγωγή τεχνολογιών όπως η υπολογιστική νέφους. Συγκεκριμένα, έχουν αναγνωριστεί οι ακόλουθοι βασικοί άξονες:

1. Μελέτη της ήδη υπάρχουσας δικτυακής υποδομής και έρευνα για νέους τρόπους διασύνδεσης των δρομολογητών, λαμβάνοντας υπόψη τα ιδιαίτερα χαρακτηριστικά των πληροφοριοκεντρικών δικτύων. Εργασίες στον τομέα της βελτιστοποίησης, όπως η [111], δύνανται να επεκταθούν προκειμένου να ενσωματώνουν στη διατύπωση του προβλήματος τη διαφοροποίηση που προκαλεί στην κίνηση του συστήματος η κρυφή μνήμη. Τα αποτελέσματα μπορούν να φανούν ιδιαίτερα χρήσιμα στους παρόχους δικτύου, προκειμένου να προχωρήσουν στο μελλοντικό σχεδιασμό των υποδομών τους.
2. Ικανότητα κλιμάκωσης των πινάκων δρομολόγησης των πληροφοριοκεντρικών δικτύων. Με δεδομένο ότι το εν λόγω μοντέλο δρομολογεί το περιεχόμενο με βάση το όνομα, υπάρχουν προφανή προβλήματα στην κατασκευή και διατήρηση των πινάκων δρομολόγησης. Το πρόβλημα αυτό δε, γίνεται ακόμα πιο δυσεπίλυτο λαμβάνοντας υπόψη ότι η ανεύρεση του επόμενου κόμβου όπου θα προωθηθεί το εκάστοτε πακέτο, θα πρέπει να γίνεται σε χρόνο γραμμής (line speed) προκειμένου να μην εισάγει το δίκτυο καθυστερήσεις στη μετάδοση. Συνεπώς, οι προτεινόμενες λύσεις θα πρέπει να εκτείνονται τόσο στο αλγοριθμικό κομμάτι της διατήρησης των πινάκων δρομολόγησης και της αναζήτησης με αποδοτικό τρόπο αλλά και και στο κομμάτι του υλικού, αξιολογώντας τις δυνατότητες των διαφόρων τεχνολογιών αποθηκευτικών μέσων, τόσο για μόνιμων όσο και προσωρινών (π.χ. SRAM, DRAM, SSD κλπ).
3. Υλοποίηση των προτεινόμενων επεκτάσεων του CCNx και δοκιμές σε πραγματικό περιβάλλον εξομοίωσης (testbed). Είναι γεγονός ότι οι προσομοιώσεις, αν και μπορούν να δώσουν μια πολύ καλή εκτίμηση της απόδοσης ενός πρωτοκόλλου, είναι

πολύ δύσκολο να αποτυπώσουν πλήρως το πραγματικό περιβάλλον λειτουργίας που υπάρχει αυτή τη στιγμή. Για το λόγο αυτό έχει προχωρήσει πολύ η έρευνα πάνω στη δημιουργία και διαχείριση υποδομών, οι οποίες επιτρέπουν τη διεξαγωγή πειραμάτων σε ένα ρεαλιστικό περιβάλλον, με πιο γνωστή ίσως το PlanetLab. Δεδομένου ότι ο πηγαίος κώδικας του CCNx είναι ανοικτός, έχει ιδιαίτερο ενδιαφέρον να παρατηρηθεί κατά πόσο τα αποτελέσματα των προσομοιώσεων επιβεβαιώνονται και σε ένα τέτοιο περιβάλλον εξομοίωσης.

4. Δημιουργία ενός ολοκληρωμένου περιβάλλοντος προσομοίωσης του CCNx με δυνατότητα παράλληλης εκτέλεσης. Ο προσομοιωτής JCCNxSim αποτελεί το πρώτο βήμα προς αυτή την κατεύθυνση με την ενσωμάτωση των μηχανισμών του CCNx. Ωστόσο, μια μεγάλη συνεισφορά στην κοινότητα των πληροφοριοκεντρικών δικτύων θα ήταν η δημιουργία μιας υπηρεσίας, όπου οι ερευνητές θα μπορούν να υποβάλουν τα πειράματά τους, δοκιμάζοντας διαφορετικά πρωτόκολλα και υποδομές. Μια τέτοια υπηρεσία θα ήταν ιδιαίτερα χρήσιμο να παρέχει δυνατότητα παρακολούθησης της εξέλιξης του πειράματος σε πραγματικό χρόνο αλλά και διατήρηση των δεδομένων σε βάσεις δεδομένων, ώστε τα τελευταία να είναι διαθέσιμα και στο μέλλον στους ερευνητές. Επιπλέον, μια ιδιαίτερα χρήσιμη προσέγγιση για μια τέτοια υπηρεσία είναι να επιτρέψει σε ερευνητές να διαμοιράζονται μεταξύ τους τα αποτελέσματά τους προκειμένου να καθίστανται πιο εύκολες οι συγκρίσεις μέσα, μάλιστα, από το ίδιο περιβάλλον με τις πραγματικές υλοποιήσεις του καθενός.
5. Στατιστική μελέτη της επαναληπτικότητας των ερωτημάτων σε κατανεμημένο σύστημα αναζήτησης. Αν και έχει γίνει αναλυτική έρευνα ως προς τα ενδιαφέροντα των χρηστών γενικά, με βάση τον αριθμό των αιτήσεων σε ιστοσελίδες [27] ή βίντεο [62] [112], η επαναληπτικότητα των ερωτημάτων στις μηχανές αναζήτησης έχει μελετηθεί είτε παλιότερα [41] είτε σε περιορισμένο σύνολο δεδομένων [113]. Δεδομένων των στοιχείων που ήδη παρέχει για παράδειγμα η Google [61] [114], μπορεί να γίνει αναλυτική στατιστική μελέτη, ώστε να εξαχθούν χρήσιμα συμπεράσματα ως προς την επαναληπτικότητα των ερωτημάτων με βάση διάφορα κοινωνικά χαρακτηριστικά, πράγμα που μπορεί να οδηγήσει στο σχεδιασμό νέων αντίστοιχων πρωτοκόλλων αναζήτησης και δημοσίευσης/συνδρομής, επικεντρωμένων σε αυτά τα δεδομένα.
6. Ενσωμάτωση τεχνολογιών υπολογιστής νέφους (cloud computing) στο κομμάτι της οργάνωσης του περιεχομένου. Συγκεκριμένα, τόσο η υλοποιήσεις των αντίστοιχων μηχανών του MPEG-M (π.χ. OverlayEngine) όσο και του κατανεμημένου συστήματος αρχείων EuterpeFS μπορούν να αναθεωρηθούν για να λάβουν υπόψη τους τα πλεονεκτήματα που φέρνει στο φως η υπολογιστική νέφους. Συστήματα αρχείων όπως το Apache Hadoop [115] φέρνουν μια νέα λογική στη διαχείριση πόρων σε τέτοια περιβάλλοντα, συνεπώς η ενοποίησή τους με το EuterpeFS μπορεί να επιφέρει επιπλέον πλεονεκτήματα στους (ολοένα αυξανόμενους) χρήστες του νέφους.

Βιβλιογραφία

- [1] “Scalable and Adaptive Internet Solutions (SAIL).” <http://www.sail-project.eu/>, Last accessed: 14 August 2013.
- [2] “Named Data Networking (NDN).” <http://www.named-data.net/>, Last accessed: 14 August 2013.
- [3] “PURSUIT.” <http://www.fp7-pursuit.eu/PursuitWeb/>, Last accessed: 14 August 2013.
- [4] “Content Mediator architecture for content-aware nETworks (COMET).” <http://www.comet-project.org/>, Last accessed: 14 August 2013.
- [5] “The CONVERGENCE Project.” <http://www.ict-convergence.eu/>, Last accessed: 14 August 2013.
- [6] P. Gkonis, C. Patrikakis, A. Anadiotis, D. Kaklamani, M. Andrade, A. Detti, G. Tropea, and N. Melazzi, “A content-centric, publish-subscribe architecture delivering mobile context-aware health services,” in *Future Network Mobile Summit (FutureNetw)*, 2011, pp. 1–9, June 2011.
- [7] L. Chiariglione, A. Difino, N. B. Melazzi, S. Salsano, A. Detti, G. Tropea, A. C. G. Anadiotis, A. S. Mousas, I. S. Venieris, and C. Z. Patrikakis, “Publish/subscribe over information centric networks: a standardized approach in convergence (accepted for publication),” in *Future Network Mobile Summit (FutureNetw)*, 2012.
- [8] G. Lioudakis, A. Anadiotis, A. Mousas, C. Patrikakis, D. Kaklamani, and I. Venieris, “Routing in content-centric networks: From names to concepts,” in *New Technologies, Mobility and Security (NTMS), 2012 5th International Conference on*, pp. 1–5, May 2012.
- [9] “Content-Oriented Networking: a New Experience for Content Transfer (ANR-CONNECT).” <http://www.anr-connect.org/>, Last accessed: 14 August 2013.
- [10] “Information-Centric Networking Research Group (ICNRG).” <https://trac.tools.ietf.org/group/irtf/trac/wiki/icnrg>, Last accessed: 14 August 2013.
- [11] “Project CCNx.” <http://www.ccnx.org/>, Last accessed: 14 August 2013.
- [12] A. Ghodsi, T. Koponen, J. Rajahalme, P. Sarolahti, and S. Shenker, “Naming in content-oriented architectures,” in *Proceedings of the ACM SIGCOMM workshop on Information-centric networking, ICN '11*, (New York, NY, USA), pp. 1–6, ACM, 2011.
- [13] M. Gritter and D. R. Cheriton, “An architecture for content routing support in the internet,” in *Proceedings of the 3rd conference on USENIX Symposium on Internet Technologies and Systems - Volume 3*, USITS'01, (Berkeley, CA, USA), pp. 4–4, USENIX Association, 2001.
- [14] Y. Rekhter, T. Li, and S. Hares, “A Border Gateway Protocol 4 (BGP-4),” Request for Comments: 4271, Network Working Group, January 2006.
- [15] T. Koponen, M. Chawla, B.-G. Chun, A. Ermolinskiy, K. H. Kim, S. Shenker, and I. Stoica, “A data-oriented (and beyond) network architecture,” *SIGCOMM Comput. Commun. Rev.*, vol. 37, pp. 181–192, Aug. 2007.
- [16] P. T. Eugster, P. A. Felber, R. Guerraoui, and A.-M. Kermarrec, “The many faces of publish/subscribe,” *ACM Comput. Surv.*, vol. 35, pp. 114–131, June 2003.

- [17] I. Stoica, D. Adkins, S. Zhuang, S. Shenker, and S. Surana, "Internet indirection infrastructure," *IEEE/ACM Trans. Netw.*, vol. 12, pp. 205–218, Apr. 2004.
- [18] I. Stoica, R. Morris, D. Liben-Nowell, D. R. Karger, M. F. Kaashoek, F. Dabek, and H. Balakrishnan, "Chord: a scalable peer-to-peer lookup protocol for internet applications," *IEEE/ACM Trans. Netw.*, vol. 11, pp. 17–32, Feb. 2003.
- [19] V. Jacobson, D. K. Smetters, J. D. Thornton, M. F. Plass, N. H. Briggs, and R. L. Braynard, "Networking named content," in *Proceedings of the 5th international conference on Emerging networking experiments and technologies*, CoNEXT '09, (New York, NY, USA), pp. 1–12, ACM, 2009.
- [20] A. Detti, N. Blefari Melazzi, S. Salsano, and M. Pomposini, "Conet: a content centric inter-networking architecture," in *Proceedings of the ACM SIGCOMM workshop on Information-centric networking*, ICN '11, (New York, NY, USA), pp. 50–55, ACM, 2011.
- [21] "PeerSim: A Peer-to-Peer Simulator." <http://peersim.sourceforge.net/>, Last accessed: 14 August 2013.
- [22] A. Montresor and M. Jelasity, "PeerSim: A scalable P2P simulator," in *Proc. of the 9th Int. Conference on Peer-to-Peer (P2P'09)*, 2009.
- [23] A. Detti, M. Pomposini, N. Blefari-Melazzi, and S. Salsano, "Supporting the web with an information centric network that routes by name," *Computer Networks*, vol. 56, no. 17, pp. 3705–3722, 2012.
- [24] Sandvine Intelligent Broadband Networks, "Global Internet Phenomena Report, 1H 2013," White Paper, November 2013.
- [25] B. Cohen, "The BitTorrent Protocol Specification," Standard, BitTorrent.org, January 2008.
- [26] "BitTorrent.org." [Online]. http://www.bittorrent.org/beps/bep_0003.html. Last accessed: 15 January 2013.
- [27] L. Breslau, P. Cao, L. Fan, G. Phillips, and S. Shenker, "Web caching and zipf-like distributions: evidence and implications," in *INFOCOM '99. Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 1, pp. 126–134 vol.1, 1999.
- [28] CISCO, "Cisco Virtual Networking Index: Forecast and Methodology, 2012-2017," White Paper, May 2013.
- [29] R. Pantos and W. May, "HTTP Live Streaming," tech. rep., October 2012.
- [30] A. Zambelli, "IIS Smooth Streaming Technical Overview," tech. rep., March 2009.
- [31] L. Larson-Kelley, "Introducing Adobe Media Server 5," tech. rep., 2012.
- [32] I. Sodagar, "The mpeg-dash standard for multimedia streaming over the internet," *IEEE Multimedia*, vol. 18, no. 4, pp. 62–67, 2011.
- [33] D. Kulinski and J. Burke, "NDN Video: Random-access Live and Pre-recorded Streaming using NDN," tech. rep., 9 2012.
- [34] A. Detti, M. Pomposini, N. Blefari-Melazzi, S. Salsano, and A. Bragagnini, "Offloading cellular networks with Information-Centric Networking: The case of video streaming.," in *WOWMOM*, pp. 1–3, IEEE Computer Society, 2012.
- [35] W. So, A. Narayanan, D. Oran, and Y. Wang, "Toward fast ndn software forwarding lookup engine based on hash tables," in *Proceedings of the eighth ACM/IEEE symposium on Architectures for networking and communications systems*, ANCS '12, (New York, NY, USA), pp. 85–86, ACM, 2012.
- [36] K. Shanmugasundaram, H. Brönnimann, and N. Memon, "Payload attribution via hierarchical bloom filters," in *Proceedings of the 11th ACM conference on Computer and communications security*, CCS '04, (New York, NY, USA), pp. 31–41, ACM, 2004.
- [37] L. Muscariello, G. Carofiglio, and M. Gallo, "Bandwidth and storage sharing performance in information centric networking," in *Proceedings of the ACM SIGCOMM workshop on Information-centric networking*, ICN '11, (New York, NY, USA), pp. 26–31, ACM, 2011.

- [38] C. Liu, I. Bouazizi, and M. Gabbouj, "Rate adaptation for adaptive http streaming," in *Proceedings of the second annual ACM conference on Multimedia systems, MMSys '11*, (New York, NY, USA), pp. 169–174, ACM, 2011.
- [39] B. Wang, J. Kurose, P. Shenoy, and D. Towsley, "Multimedia streaming via tcp: An analytic performance study," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 4, pp. 16:1–16:22, May 2008.
- [40] A.-C. G. Anadiotis, C. Z. Patrikakis, and I. S. Venieris, "On the performance improvement of gossip protocols for content-based publish-subscribe through caching," *Computer Networks*, vol. 57, no. 18, pp. 3759 – 3772, 2013.
- [41] B. Smyth, E. Balfe, J. Freyne, P. Briggs, M. Coyle, and O. Boydell, "Exploiting query repetition and regularity in an adaptive community-based web search engine," *User Modeling and User-Adapted Interaction*, vol. 14, pp. 383–423, Jan. 2005.
- [42] S. Tang, E. Jaho, I. Stavrakakis, I. Koukoutsidis, and P. V. Mieghem, "Modeling gossip-based content dissemination and search in distributed networking," *Computer Communications*, vol. 34, no. 6, pp. 765 – 779, 2011.
- [43] P. T. Eugster, R. Guerraoui, S. B. Handurukande, P. Kouznetsov, and A.-M. Kermarrec, "Lightweight probabilistic broadcast," *ACM Trans. Comput. Syst.*, vol. 21, pp. 341–374, Nov. 2003.
- [44] A.-M. Kermarrec, L. Massoulié, and A. Ganesh, "Probabilistic reliable dissemination in large-scale systems," *Parallel and Distributed Systems, IEEE Transactions on*, vol. 14, no. 3, pp. 248–258, 2003.
- [45] S. Ratnasamy, P. Francis, M. Handley, R. Karp, and S. Shenker, "A scalable content-addressable network," in *Proceedings of the 2001 conference on Applications, technologies, architectures, and protocols for computer communications, SIGCOMM '01*, (New York, NY, USA), pp. 161–172, ACM, 2001.
- [46] A. I. T. Rowstron and P. Druschel, "Pastry: Scalable, decentralized object location, and routing for large-scale peer-to-peer systems," in *Proceedings of the IFIP/ACM International Conference on Distributed Systems Platforms Heidelberg, Middleware '01*, (London, UK, UK), pp. 329–350, Springer-Verlag, 2001.
- [47] P. Maymounkov and D. Mazières, "Kademlia: A peer-to-peer information system based on the xor metric," in *Revised Papers from the First International Workshop on Peer-to-Peer Systems, IPTPS '01*, (London, UK, UK), pp. 53–65, Springer-Verlag, 2002.
- [48] R. Ahmed and R. Boutaba, "A survey of distributed search techniques in large scale distributed systems," *Communications Surveys Tutorials, IEEE*, vol. 13, pp. 150 –167, quarter 2011.
- [49] S. Ferretti, "Publish-subscribe systems via gossip: a study based on complex networks," in *Proceedings of the Fourth Annual Workshop on Simplifying Complex Networks for Practitioners, SIMPLEX '12*, (New York, NY, USA), pp. 7–12, ACM, 2012.
- [50] P. T. Eugster, R. Guerraoui, A.-M. Kermarrec, and L. Massoulié, "Epidemic information dissemination in distributed systems," *Computer*, vol. 37, pp. 60–67, May 2004.
- [51] E. Cohen, A. Fiat, and H. Kaplan, "Associative search in peer to peer networks: Harnessing latent semantics," *Comput. Netw.*, vol. 51, pp. 1861–1881, June 2007.
- [52] H. Cai and J. Wang, "Foreseer: a novel, locality-aware peer-to-peer system architecture for keyword searches," in *Proceedings of the 5th ACM/IFIP/USENIX international conference on Middleware, Middleware '04*, (New York, NY, USA), pp. 38–58, Springer-Verlag New York, Inc., 2004.
- [53] G. Chockler, R. Melamed, Y. Tock, and R. Vitenberg, "Spidercast: a scalable interest-aware overlay for topic-based pub/sub communication," in *Proceedings of the 2007 inaugural international conference on Distributed event-based systems, DEBS '07*, (New York, NY, USA), pp. 14–25, ACM, 2007.
- [54] F. Rahimian, S. Girdzijauskas, A. H. Payberah, and S. Haridi, "Vitis: A gossip-based hybrid overlay for internet-scale publish/subscribe enabling rendezvous routing in unstructured overlay networks," in *Proceedings of the 2011 IEEE International Parallel & Distributed Processing Symposium, IPDPS '11*, (Washington, DC, USA), pp. 746–757, IEEE Computer Society, 2011.

- [55] M. Jelasity, A. Montresor, and O. Babaoglu, "T-man: Gossip-based fast overlay topology construction," *Comput. Netw.*, vol. 53, pp. 2321–2339, Aug. 2009.
- [56] M. Matos, A. Nunes, R. Oliveira, and J. Pereira, "Stan: exploiting shared interests without disclosing them in gossip-based publish/subscribe," in *Proceedings of the 9th international conference on Peer-to-peer systems, IPTPS'10*, (Berkeley, CA, USA), pp. 9–9, USENIX Association, 2010.
- [57] L. Sharifi and S. Khorsandi, "A popularity-based query scheme in p2p networks using adaptive gossip sampling," *Peer-to-Peer Networking and Applications*, vol. 6, no. 1, pp. 75–85, 2013.
- [58] Y.-H. Tan, K. Lü, and Y.-P. Lin, "Organisation and management of shared documents in super-peer networks based semantic hierarchical cluster trees," *Peer-to-Peer Networking and Applications*, vol. 5, pp. 292–308, 2012.
- [59] K. Sripanidkulchai, B. Maggs, and H. Zhang, "Efficient content location using interest-based locality in peer-to-peer systems," in *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, vol. 3, pp. 2166 – 2176 vol.3, march-3 april 2003.
- [60] E. Casalicchio and F. Morabito, "Distributed subscriptions clustering with limited knowledge sharing for content-based publish/subscribe systems," in *Network Computing and Applications, 2007. NCA 2007. Sixth IEEE International Symposium on*, pp. 105 –112, july 2007.
- [61] "Google Trends - Hot Searches." [Online]. <http://www.google.com/trends>. Last accessed: 13/11/2013.
- [62] A. Mahanti, N. Carlsson, A. Mahanti, M. Arlitt, and C. Williamson, "A tale of the tails: Power-laws in internet measurements," *Network, IEEE*, vol. 27, no. 1, pp. 59–64, 2013.
- [63] D. Karger, E. Lehman, T. Leighton, R. Panigrahy, M. Levine, and D. Lewin, "Consistent hashing and random trees: distributed caching protocols for relieving hot spots on the world wide web," in *Proceedings of the twenty-ninth annual ACM symposium on Theory of computing, STOC '97*, (New York, NY, USA), pp. 654–663, ACM, 1997.
- [64] Y. Fernandess, A. Fernández, and M. Monod, "A generic theoretical framework for modeling gossip-based algorithms," *SIGOPS Oper. Syst. Rev.*, vol. 41, pp. 19–27, Oct. 2007.
- [65] P. Eugster, P. Felber, and F. Le Fessant, "The "art" of programming gossip-based systems," *SIGOPS Oper. Syst. Rev.*, vol. 41, pp. 37–42, Oct. 2007.
- [66] "BitTorrentSpecification." [Online]. <http://wiki.theory.org/BitTorrentSpecification>. Last accessed: 15 January 2013.
- [67] International Standards Organization, "ISO/IEC TR 21000-1:2004 Information technology – Multimedia framework (MPEG-21) – Part 1: Vision, Technologies and Strategy," November 2004.
- [68] International Standards Organization, "ISO/IEC 21000-2:2005 Information technology – Multimedia framework (MPEG-21) – Part 2: Digital Item Declaration," October 2005.
- [69] ISO/IEC 21000-2:2005, "Information Technology – Multimedia Framework (MPEG-21) – Part 2: Digital Item Declaration - Schema for Derived DIDL Types."
- [70] International Standards Organization, "ISO/IEC 21000-5:2004 Information technology – Multimedia framework (MPEG-21) – Part 5: Rights Expression Language," April 2004.
- [71] International Standards Organization, "ISO/IEC 21000-6:2004 Information technology – Multimedia framework (MPEG-21) – Part 6: Rights Data Dictionary," May 2004.
- [72] P. Kudumakis, M. Sandler, A.-C. G. Anadiotis, I. S. Venieris, A. Difino, X. Wang, G. Tropea, M. Grafl, V. Rodríguez-Doncel, S. Llorente, and J. Delgado, "Mpeg-m: A digital media ecosystem for interoperable applications," *Signal Processing: Image Communication*, no. 0, pp. –, 2013.
- [73] A.-C. Anadiotis, G. Tropea, and I. Venieris, "Input contribution on overlay technology," November 2011. Input document to the International Standards Organization, ISO/IEC JTC 1/SC 29/WG 11 (MPEG).

- [74] A. Difino, A. Mousas, A.-C. Anadiotis, B. Ardeleanu, and P. Gkonis, "Proposed revised version of MPEG-M part2," April 2012. Input document to the International Standards Organization, ISO/IEC JTC 1/SC 29/WG 11 (MPEG).
- [75] A.-C. Anadiotis, G. Tropea, and I. Venieris, "Proposal for adding the overlay technology engine," April 2012. Input document to the International Standards Organization, ISO/IEC JTC 1/SC 29/WG 11 (MPEG).
- [76] A. Difino, A. Mousas, and A.-C. Anadiotis, "Proposed revised version of mpeg-m part2," June 2012. Input document to the International Standards Organization, ISO/IEC JTC 1/SC 29/WG 11 (MPEG).
- [77] A. Difino, A.-C. Anadiotis, and G. Tropea, "Proposal for reengineering of MPEG-M reference software," July 2011. Input document to the International Standards Organization, ISO/IEC JTC 1/SC 29/WG 11 (MPEG).
- [78] A. Difino, A.-C. Anadiotis, and I. Venieris, "Proposed pe/es implementation of mpeg-m part3," November 2011. Input document to the International Standards Organization, ISO/IEC JTC 1/SC 29/WG 11 (MPEG).
- [79] A.-C. Anadiotis, A. Difino, A. Mousas, S. Signorello, and G. Tropea, "An mpeg-m use case: The mpm photo sharing service," February 2012. Input document to the International Standards Organization, ISO/IEC JTC 1/SC 29/WG 11 (MPEG).
- [80] A. Difino, A. Mousas, A.-C. Anadiotis, B. Ardeleanu, and P. Gkonis, "Proposed revised version of MPEG-M part3," April 2012. Input document to the International Standards Organization, ISO/IEC JTC 1/SC 29/WG 11 (MPEG).
- [81] A. Difino, G. Tropea, A.-C. Anadiotis, and S. Llorente, "MPEG-M reference software workplan," April 2012. Input document to the International Standards Organization, ISO/IEC JTC 1/SC 29/WG 11 (MPEG).
- [82] A. Difino, A. Mousas, A.-C. Anadiotis, B. Ardeleanu, and P. Gkonis, "Proposal for a refactored version of mxm for mpeg-m purpose," June 2012. Input document to the International Standards Organization, ISO/IEC JTC 1/SC 29/WG 11 (MPEG).
- [83] A. Difino, A. Mousas, and A.-C. Anadiotis, "Proposed revised version of mpeg-m part3," June 2012. Input document to the International Standards Organization, ISO/IEC JTC 1/SC 29/WG 11 (MPEG).
- [84] A.-C. Anadiotis and L. Chiariglione, "Introducing user identification in mpeg-m," November 2011. Input document to the International Standards Organization, ISO/IEC JTC 1/SC 29/WG 11 (MPEG).
- [85] A.-C. Anadiotis, A. Difino, and G. Tropea, "Proposal for post content elementary service," February 2012. Input document to the International Standards Organization, ISO/IEC JTC 1/SC 29/WG 11 (MPEG).
- [86] A.-C. Anadiotis, A. Difino, and I. Venieris, "First step towards unified orchestrator/aggregation environment," November 2011. Input document to the International Standards Organization, ISO/IEC JTC 1/SC 29/WG 11 (MPEG).
- [87] ISO/IEC 23006-1:2013, "Information technology – MPEG extensible middleware (MXM) – Part 1: Architecture."
- [88] ISO/IEC 23006-2:2013, "Information technology – MPEG extensible middleware (MXM) – Part 2: MPEG extensible middleware (MXM) API."
- [89] ISO/IEC 23006-3:2013, "Information technology – MPEG extensible middleware (MXM) – Part 3: Conformance and reference software."
- [90] ISO/IEC 23006-4:2013, "Information technology – MPEG extensible middleware (MXM) – Part 4: Elementary Services."
- [91] Object Management Group (OMG), "Business Process Model and Notation (BPMN) Version 2.0." [Online]. <http://www.omg.org/spec/BPMN/2.0/>. Last accessed: 23/11/2013, January 2011.

- [92] ISO/IEC 23006-5:2013, "Information technology – Multimedia service platform technologies – Part 5: Service aggregation."
- [93] A.-C. Anadiotis and I. S. Venieris, "Distributed Content Semantic Organization and Access Control with EuterpeFS," *Journal of Internet Technology*, vol. 14, pp. 643–650, July 2013.
- [94] G. F. Coulouris, J. Dollimore, and T. Kindberg, *Distributed systems: concepts and design*. Essex, England: Pearson Education Ltd, 2001.
- [95] A. S. Tanenbaum, *Σύγχρονα Λειτουργικά Συστήματα - Τόμος Β' (Κατανεμημένα Συστήματα)*. Prentice Hall International Inc., Α. Παπασωτηρίου και ΣΙΑ Ο.Ε., 1994.
- [96] "stat(2): file status - Linux man page."
- [97] "The Linux Kernel - Chapter 9 The File System."
- [98] A. S. W. A. S. Tanenbaum, *Operating Systems, Design and Implementation, The MINIX Book*. Pearson International Edition, 2009.
- [99] "Sourceforge.net: FileSystems - fuse." [Online]. <http://sourceforge.net/apps/mediawiki/fuse/index.php?title=FileSystems>. Last accessed: 19/11/2013.
- [100] Network Working Group, Sun Microsystems Inc., "NFS: Network File System Protocol Specification," tech. rep., March 1989. Request for Comments: 1094.
- [101] R. Sandberg, D. Golgberg, S. Kleiman, D. Walsh, and B. Lyon, "Innovations in internetworking," ch. Design and implementation of the Sun network filesystem, pp. 379–390, Norwood, MA, USA: Artech House, Inc., 1988.
- [102] S. Shepler, B. Callaghan, D. Robinson, R. Thurlow, C. Beame, M. Eisler, D. Noveck, "Network File System (NFS) version 4 Protocol." RFC 3530, April 2003.
- [103] Sun Microsystems Inc., *NFS Administration Guide*, August 1997.
- [104] J. H. Howard, M. L. Kazar, S. G. Menees, D. A. Nichols, M. Satyanarayanan, R. N. Sidebotham, and M. J. West, "Scale and performance in a distributed file system," *ACM Trans. Comput. Syst.*, vol. 6, no. 1, pp. 51–81, 1988.
- [105] M. Satyanarayanan, "The evolution of coda," *ACM Trans. Comput. Syst.*, vol. 20, no. 2, pp. 85–124, 2002.
- [106] Sun Microsystems, "LUSTRE FILE SYSTEM. High Performance Storage Architecture and Scalable Cluster File System," 2007.
- [107] Sun Microsystems, "Lustre HPCS Design Overview," 2009.
- [108] A. Mousas, A.-C. Anadiotis, G. Lioudakis, J. Papanis, P. Gkonis, D. Kaklamani, and I. Venieris, "On supporting secure information distribution in heterogeneous systems using standard technologies," *Wireless Personal Communications*, pp. 1–21, 2013.
- [109] J. Alpert and N. Hajaj, "We knew the web was big...." [Online]. <http://googleblog.blogspot.gr/2008/07/we-knew-web-was-big.html>. Last accessed: 7/12/2013, 2008.
- [110] A. Ghodsi, S. Shenker, T. Koponen, A. Singla, B. Raghavan, and J. Wilcox, "Information-centric networking: Seeing the forest for the trees," in *Proceedings of the 10th ACM Workshop on Hot Topics in Networks, HotNets-X*, (New York, NY, USA), pp. 1:1–1:6, ACM, 2011.
- [111] C. Pappas, A.-C. Anadiotis, C. Papagianni, and I. Venieris, "Heuristics for the multi-level capacitated minimum spanning tree problem," *Optimization Letters*, pp. 1–12, 2013.
- [112] M. Cha, H. Kwak, P. Rodriguez, Y.-Y. Ahn, and S. Moon, "I tube, you tube, everybody tubes: Analyzing the world's largest user generated content video system," in *Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement, IMC '07*, (New York, NY, USA), pp. 1–14, ACM, 2007.

- [113] C. Bauckhage, "Insights into internet memes.," in *ICWSM* (L. A. Adamic, R. A. Baeza-Yates, and S. Counts, eds.), The AAAI Press, 2011.
- [114] "Google zeitgeist." [Online]. <http://www.google.com/zeitgeist/>. Last accessed: 7/12/2013.
- [115] "Welcome to apache hadoop." [Online]. <http://hadoop.apache.org/>. Last accessed: 7/12/2013.

Δημοσιεύσεις

Άρθρα σε Επιστημονικά Περιοδικά

- [1] A.-C. G. Anadiotis, C. Z. Patrikakis, and I. S. Venieris, "On the performance improvement of gossip protocols for content-based publish-subscribe through caching," *Computer Networks*, vol. 57, no. 18, pp. 3759 – 3772, 2013.
- [2] P. Kudumakis, M. Sandler, A.-C. G. Anadiotis, I. S. Venieris, A. Difino, X. Wang, G. Tropea, M. Grafl, V. Rodríguez-Doncel, S. Llorente, and J. Delgado, "Mpeg-m: A digital media ecosystem for interoperable applications," *Signal Processing: Image Communication*, no. 0, pp. –, 2013.
- [3] A.-C. Anadiotis and I. S. Venieris, "Distributed Content Semantic Organization and Access Control with EuterpeFS," *Journal of Internet Technology*, vol. 14, pp. 643–650, July 2013.
- [4] A. Mousas, A.-C. Anadiotis, G. Lioudakis, J. Papanis, P. Gkonis, D. Kaklamani, and I. Venieris, "On supporting secure information distribution in heterogeneous systems using standard technologies," *Wireless Personal Communications*, pp. 1–21, 2013.
- [5] C. Pappas, A.-C. Anadiotis, C. Papagianni, and I. Venieris, "Heuristics for the multi-level capacitated minimum spanning tree problem," *Optimization Letters*, pp. 1–12, 2013.
- [6] F. Almeida, H. Castro, M.-T. Andrade, G. Tropea, S. Signorello, N. Blefari Melazzi, A. S. Mousas, A.-C. G. Anadiotis, D. I. Kaklamani, I. S. Venieris, S. Minelli, and A. Difino, "Digital forgetting in information centric networks - the convergence perspective," accepted for publication in *New Review of Hypermedia and Multimedia*.
- [7] A.-C. G. Anadiotis, P. K. Gkonis, D. I. Kaklamani, and I. S. Venieris, "A cross-layer information centric-relay transmission strategy for mimo-ofdma multicellular networks," accepted for publication in *Wireless Communications and Mobile Computing*.

Κεφάλαια σε Βιβλία

- [1] G. Tropea, G. Bianchi, N. Blefari Melazzi, H. Castro, L. Chiariglione, A. Difino, T. Huebner, A.-C. Anadiotis, and A. Mousas, "The Adoption of Rights Expression Language in CONVERGENCE,"

in *Enhancing the Internet with the CONVERGENCE System* (F. Almeida, M. T. Andrade, N. Blefari Melazzi, R. Walker, H. Hussmann, and I. S. Venieris, eds.), Signals and Communication Technology, Springer London, 2014.

- [2] A.-C. Anadiotis, A. Mousas, A. Difino, and C. Patrikakis, “The content level (comid),” in *Enhancing the Internet with the CONVERGENCE System* (F. Almeida, M. T. Andrade, N. Blefari Melazzi, R. Walker, H. Hussmann, and I. S. Venieris, eds.), Signals and Communication Technology, pp. 73–102, Springer London, 2014.

Πρακτικά Συνεδρίων

- [1] D. Syrivelis, A.-C. Anadiotis, A. Apostolaras, T. Korakis, and L. Tassiulas, “Tlqap: A topology and link quality assessment protocol for efficient node allocation on wireless testbeds,” in *Proceedings of the 4th ACM International Workshop on Experimental Evaluation and Characterization, WINTECH '09*, (New York, NY, USA), pp. 27–34, ACM, 2009.
- [2] A.-C. Anadiotis, A. Apostolaras, D. Syrivelis, T. Korakis, L. Tassiulas, L. Rodriguez, and M. Ott, “A new slicing scheme for efficient use of wireless testbeds,” in *Proceedings of the 4th ACM International Workshop on Experimental Evaluation and Characterization, WINTECH '09*, (New York, NY, USA), pp. 83–84, ACM, 2009.
- [3] A.-C. Anadiotis, A. Apostolaras, D. Syrivelis, T. Korakis, L. Tassiulas, L. Rodriguez, I. Seskar, and M. Ott, “Towards maximizing wireless testbed utilization using spectrum slicing,” in *Testbeds and Research Infrastructures. Development of Networks and Communities* (T. Magedanz, A. Gavras, N. Thanh, and J. Chase, eds.), vol. 46 of *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, pp. 299–314, Springer Berlin Heidelberg, 2011.
- [4] C. Patrikakis, A.-C. Anadiotis, P. Santi, and N. Blefari-Melazzi, “Designing and experimenting a hybrid social network made up of people, agents and sensors,” in *Global Telecommunications Conference (GLOBECOM 2011), 2011 IEEE*, pp. 1–5, 2011.
- [5] P. Gkonis, C. Patrikakis, A. Anadiotis, D. Kaklamani, M. Andrade, A. Detti, G. Tropea, and N. Melazzi, “A content-centric, publish-subscribe architecture delivering mobile context-aware health services,” in *Future Network Mobile Summit (FutureNetw), 2011*, pp. 1–9, june 2011.
- [6] N. Melazzi, S. Salsano, A. Detti, G. Tropea, L. Chiariglione, A. Difino, A. Anadiotis, A. Mousas, I. Venieris, and C. Patrikakis, “Publish/subscribe over information centric networks: A standardized approach in convergence,” in *Future Network Mobile Summit (FutureNetw), 2012*, pp. 1–8, 2012.

- [7] G. Lioudakis, A. Anadiotis, A. Mousas, C. Patrikakis, D. Kaklamani, and I. Venieris, "Routing in content-centric networks: From names to concepts," in *New Technologies, Mobility and Security (NTMS), 2012 5th International Conference on*, pp. 1–5, may 2012.
- [8] F. I. Gogoulos, D. I. Kaklamani, I. S. Venieris, A.-C. Anadiotis, U. Inden, S. Naimark, P. Skobelev, and A. Tsarev, "An Intelligent Tracking Operations Management System," in *Proceedings of the 1st IFAC Workshop on Advances in Control and Automation Theory for Transportation Applications (ACATTA)*, 9 2013.
- [9] U. Inden, D. Meridou, M.-E. Papadopoulou, A.-C. Anadiotis, and C.-P. Rückemann, "Complex Landscapes of Risk in Operations Systems Aspects of Modelling and Processing," in *The Third International Conference on Advanced Communications and Computation (INFOCOMP)*, 11 2013.

