ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΕΠΙΚΟΙΝΩΝΙΩΝ, ΗΛΕΚΤΡΟΝΙΚΗΣ ΚΑΙ ΣΥΣΤΗΜΑΤΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ

Υπηρεσίες Διαχείρισης, Ελέγχου και Προώθησης Δεδομένων σε Περιβάλλοντα Εικονικής Δικτύωσης Οριζόμενης από Λογισμικό (SDN)

Management, Control and Data Plane Services in Software Defined Networking (SDN)

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Χρήστος Γ. Αργυρόπουλος

Αθήνα, Ιούνιος 2014

πᾶσά τε ἐπιστήμη χωριζομένη δικαιοσύνης καὶ τῆς ἄλλης ἀρετῆς πανουργία, οὐ σοφία φαίνεται.
ΠΛΑΤΩΝΟΣ ΜΕΝΕΞΕΝΟΣ <<246δ - 247α>>

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΕΠΙΚΟΙΝΩΝΙΩΝ, ΗΛΕΚΤΡΟΝΙΚΗΣ ΚΑΙ ΣΥΣΤΗΜΑΤΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ

Υπηρεσίες Διαχείρισης, Ελέγχου και Προώθησης Δεδομένων σε Περιβάλλοντα Εικονικής Δικτύωσης Οριζόμενης από Λογισμικό (SDN)

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Χρήστος Γ. Αργυρόπουλος

Συμβουλευτική Επιτροπή : Βασίλειος Μάγκλαρης, Καθηγητής Ε.Μ.Π
Συμεών Παπαβασιλείου, Αν. Καθηγητής Ε.Μ.Π
Δημήτριος Καλογεράς, Ερευνητής Β΄ Ε.Π.Ι.Σ.Ε.Υ

Εγκρίθηκε από την επταμελή εξεταστική επιτροπή την:

.....	.....	.....
Βασίλειος Μάγκλαρης	Συμεών Παπαβασιλείου	Δημήτριος Καλογεράς
Καθηγητής Ε.Μ.Π	Αν. Καθηγητής Ε.Μ.Π	Ερευνητής Β΄ Ε.Π.Ι.Σ.Ε.Υ
.....	.....	.....
Ευστάθιος Συκάς	Νεκτάριος Κοζύρης	Λέανδρος Τασιούλας
Καθηγητής Ε.Μ.Π	Καθηγητής Ε.Μ.Π	Καθηγητής Παν. Θεσσαλίας
	.....	
	Σπύρος Δενάζης	
	Αν. Καθηγητής Παν. Πατρών	

Αθήνα, Ιούνιος 2014

.....

Χρήστος Γ. Αργυρόπουλος

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

Περίληψη

Η παρούσα Διδακτορική Διατριβή πραγματεύεται το πρόβλημα διασύνδεσης ετερογενών υποδομών εικονικών υπολογιστικών πόρων στο επίπεδο προώθησης πακέτων και τον εμπλουτισμό των δομών του επιπέδου ελέγχου και του επιπέδου διαχείρισης σε περιβάλλοντα εικονικής δικτύωσης οριζόμενης από λογισμικό (Software Defined Networking – SDN) πολλαπλών χρηστών/ενοικιαστών.

Το σύνολο της μελέτης επικεντρώνεται σε τρεις βασικούς άξονες με κριτήριο το επίπεδο λειτουργίας δικτύου:

Επίπεδο προώθησης πακέτων (data-plane)

Δημιουργία μηχανισμού διασύνδεσης (stitching) πολλαπλών εικονικών υπολογιστικών πόρων ετερογενών ομόσπονδων υποδομών με μικρή χρονική επιβάρυνση και μικρή μείωση του ρυθμού μετάδοσης των δεδομένων. Η δυναμική απόδοση των απαιτούμενων πόρων γίνεται με κριτήριο την κλιμακοθετησιμότητα (scalability).

Επίπεδο διαχείρισης (management-plane)

Μελέτη εγγενών δυνατοτήτων/αδυναμιών του OpenFlow για παθητική παρακολούθηση και δημιουργία μηχανισμού για την αποδοτική παρακολούθηση σε περιβάλλοντα SDN υψηλού ρυθμού ροής δεδομένων, βασιζόμενο στο OpenFlow και το sFlow. Γίνεται πειραματική μελέτη των επιπτώσεων των τεχνικών δειγματοληψίας του sFlow στον εντοπισμό δικτυακών ανωμαλιών καθώς και των επιδράσεων που έχουν οι τελευταίες στο επίπεδο κεντρικοποιημένου ελέγχου δικτύων OpenFlow.

Επίπεδο ελέγχου (control-plane)

Δημιουργία μηχανισμού ανάθεσης ροών ανά εικονικό δίκτυο χρήστη σε περιβάλλοντα δικτύων μεγάλης κλίμακας κεντρικοποιημένου ελέγχου που κάνουν χρήση του πρωτοκόλλου OpenFlow. Γίνεται πειραματική σύγκριση των προτεινόμενων πολιτικών διαμοιρασμού του επιπέδου ελέγχου και απομόνωσης των εικονικών δικτύων, με κριτήριο απόδοσης την αύξηση των αποδεκτών αιτημάτων χρηστών και την κατανάλωση πόρων. Το κριτήριο απομόνωσης εξασφαλίζεται από τις προτεινόμενες μεθόδους διαμοιρασμού και τις αντίστοιχες αρχιτεκτονικές διαμοιρασμού του επιπέδου ελέγχου που προτείνονται.

Αρχικά μελετώνται τα επίπεδα λειτουργιών σε μοντέλα αναφοράς αρχιτεκτονικής δικτύου και δίνεται βάρος στο κεντρικοποιημένο επίπεδο ελέγχου, όπως αυτό έχει διαμορφωθεί με την χρήση του πρωτοκόλλου OpenFlow σε δίκτυα οριζόμενα από λογισμικό (SDN). Ακολουθεί αναφορά στην εικονικοποίηση δικτύων υπολογιστών και μελετώνται οι σχεδιαστικές ελλείψεις των πρώιμων υλοποιήσεων εικονικοποίησης καθώς και η εξέλιξή τους.

Η μελέτη και η δημιουργία πρωτότυπου μηχανισμού διασύνδεσης (stitching) πολλαπλών εικονικών υπολογιστικών πόρων ετερογενών ομόσπονδων υποδομών περιγράφεται στη συνέχεια και αποτελεί το πρώτο μέρος της συνεισφοράς. Τα αποτελέσματα των πειραματικών μετρήσεων της πρωτότυπης υλοποίησης δείχνουν την επίτευξη μικρής χρονικής επιβάρυνσης και μικρής μείωσης του ρυθμού μετάδοσης των δεδομένων στα εικονικά δίκτυα των χρηστών.

Στη συνέχεια, μελετάται η μέθοδος παθητικής παρακολούθησης με χρήση των εγγενών δυνατοτήτων του πρωτοκόλλου OpenFlow και των μειονεκτημάτων που παρουσιάζει η παρακολούθηση της δικτυακής κίνησης με την χρήση ενός πρωτοκόλλου κεντρικοποιημένου ελέγχου. Ακολουθεί η περιγραφή πλαισίου παθητικής παρακολούθησης υποδομών δικτυακής εικονικοποίησης που κάνει χρήση των πρωτοκόλλων OpenFlow και sFlow και έχει ως στόχο την ικανότητα χειρισμού μεγαλύτερου αριθμού ροών και ρυθμών μετάδοσης δεδομένων, σε σχέση με λύσεις που βασίζονται αποκλειστικά στο OpenFlow. Παρουσιάζονται πειραματικές μετρήσεις που προσφέρει ο μηχανισμός παθητικής παρακολούθησης δικτύων οριζόμενων από λογισμικό (προγραμματιζόμενα δίκτυα) με και χωρίς την συνδρομή του sFlow στον εντοπισμό και εξομάλυνση των δικτυακών ανωμαλιών. Βαρύτητα δίνεται στην αποτελεσματικότητα του sFlow (ευστοχία εντοπισμού δικτυακών ανωμαλιών, κατανάλωση πόρων του επιπέδου ελέγχου) και στην κλιμακοθετησιμότητα.

Το τελευταίο μέρος της εργασίας αφορά μηχανισμό ανάθεσης ροών ανά εικονικό δίκτυο χρήστη σε περιβάλλοντα SDN κεντρικοποιημένου ελέγχου που κάνουν χρήση του πρωτοκόλλου OpenFlow. Ιδιαίτερη μελέτη, μέσω της υλοποίησης πρωτότυπης μηχανής διαμοιρασμού του πεδίου ορισμού των ροών (flowspace) και πειραματικών μετρήσεων, γίνεται για να διαπιστωθεί πως περιβάλλοντα SDN κεντρικοποιημένου ελέγχου μεγάλης κλίμακας μπορούν να κάνουν χρήση του πρωτοκόλλου OpenFlow για την παροχή προηγμένων υπηρεσιών εικονικής δικτύωσης.

Εν κατακλείδι, αναφέρονται τα συμπεράσματα καθώς και οι επιστημονικές προεκτάσεις που αξιολογήθηκαν ως υψηλής ερευνητικής αξίας κατά τη συγγραφή της παρούσας Διδακτορικής Διατριβής.

Λέξεις Κλειδιά

Εικονική Δικτύωση, Δικτύωση Οριζόμενη από Λογισμικό, Πρωτόκολλο OpenFlow, Παθητική Παρακολούθηση, Μηχανισμός Ανάθεσης Ροών, Ετερογενείς Ομόσπονδες Υποδομές

Abstract

In this work, we explore the problem of interconnecting (stitching) the data-plane of heterogeneous virtualized infrastructures consisting of compute and network resources and we strive to enrich the control and management-plane functions in multi-tenant software defined networking (SDN) environments.

The overall study is focused on three major directions according to the network plane functionality:

Network data-plane

Creation of a network stitching mechanism among multiple virtual compute resources of heterogeneous infrastructures. This mechanism results in a small time overhead and reduction of data transmission rate. The dynamic resource allocation is based on the scalability criterion.

Network management-plane

Study of the inherent features/disadvantages of the OpenFlow protocol performance in passive monitoring and creation of a mechanism for efficient monitoring in SDN environments of high data flow rates, based on OpenFlow and sFlow. We experimentally study the sFlow sampling method implications in anomaly detection and the impact of these methods in the centralized control-plane of OpenFlow-enabled networks.

Network control-plane

Development of a flow delegation mechanism per tenant virtual network for WAN-scale, OpenFlow-enabled, network infrastructures. An experimental comparison of the proposed control-plane slicing and virtual network isolation methods is performed, according to the criteria of tenant request acceptance ratio and resource consumption. The proposed slicing methods and the corresponding control-plane slicing architectures ensure the prerequisite of tenant isolation.

At first, we study the network functionality planes, mainly focusing on logically centralized control-planes. The analysis is based on the way that this plane has been

defined within OpenFlow-enabled SDN context. After that, we discuss the concept of network virtualization and the design shortcomings of the early implementations.

The study and the creation of a prototype stitching mechanism among multiple virtual computing resources of heterogeneous network infrastructures are described hereafter and constitute the first part of our contribution. The experimental results of the prototype implementation demonstrate a quite small time overhead and reduction of data transmission rate in tenants' virtual networks.

Subsequently, passive monitoring approaches based on the inherent features of the OpenFlow protocol and the emerging disadvantages of using a centralized control-plane protocol for monitoring are pinpointed. After that, we describe a passive monitoring framework applicable to virtual network infrastructures that makes use of OpenFlow and sFlow, as an alternative. This framework aims at handling a larger number of flows and higher data transmission rates compared to the existing frameworks that are exclusively based on the OpenFlow protocol. We present experimental measurements of a passive monitoring mechanism applicable to SDN environments, with and without the use of sFlow, for anomaly detection and mitigation. We take under consideration sFlow effectiveness (anomaly detection effectiveness, consumption of control-plane resources) and scalability issues.

The last part of this work deals with the flow delegation mechanism per tenant virtual network in SDN environments with centralized OpenFlow-based control-planes. Special attention, via the implementation of a prototype flowspace segmentation engine and the corresponding experimental measurements, is paid in order to demonstrate that the OpenFlow protocol can be used in WAN scale network infrastructures for advanced network virtualization services provisioning.

Finally, we conclude and propose future work taking advantage of the scientific implications that were assessed during the preparation of this thesis.

Keywords

Network Virtualization, Software Defined Networking (SDN), OpenFlow Protocol, Passive Monitoring, Flow Allocation Mechanism, Heterogeneous Federated Infrastructures

Ευχαριστίες

Η πολύχρονη προσπάθεια για πνευματική και νοητική εξέλιξη, οριοθετημένη από τον στόχο εκπόνησης Διδακτορικής Διατριβής, αποτέλεσε βίωμα αναντικατάστατο. Συγκινήσεις, εντάσεις, τέλματα, προσπάθεια για κατανόηση και συγχρόνως ισορροπίες λεπτές που έπρεπε να κρατηθούν στην πορεία αναζήτησης. Εξέλιξη, απώλειες, κατακτήσεις, πείσμα, απογοητεύσεις, όμορφες στιγμές, κατανόηση, αλληλεπίδραση ανθρώπων και μεταβαλλόμενες συνθήκες σε καθημερινή βάση να οδηγούν λίγο πιο πέρα και συγχρόνως να κάνουν ορατές τις αιχμές των ορίων σε όλες τις διαστάσεις της ανθρώπινης ιδιοσυγκρασίας. Προσπάθεια.

Χωρίς να μπορώ να συγκρίνω την προσπάθεια με το αποτέλεσμα, χωρίς να μπορώ να βαθμονομήσω την κλίμακα συνεισφοράς, μπορώ στα σίγουρα να ευχαριστήσω ορισμένους ανθρώπους για αυτό το ταξίδι. Είναι άνθρωποι που βοήθησαν, δείχνοντας άλλο σημείο του ορίζοντα ο καθένας, να κινηθώ με κατεύθυνση διαφορετική από αυτή του ανεμόδαρτου κυματισμού.

Θα ξεκινήσω εκφράζοντας την ευγνωμοσύνη μου στον Επιβλέποντα Καθηγητή της Διδακτορικής μου Διατριβής και Διευθυντή του Εργαστηρίου Διαχείρισης και Βέλτιστου Σχεδιασμού Δικτύων Τηλεματικής (NETMODE) Βασίλη Μάγκλαρη. Με έκανε να συνειδητοποιήσω ότι σε μια χειμαζόμενη χώρα κάποιοι άνθρωποι προσπαθούν σταθερά, με όραμα, να οικοδομούν ένα πλούσιο οικοσύστημα προσφοράς στον ακαδημαϊκό χώρο και στο ευρύτερο σύνολο.

Θα ήθελα να ευχαριστήσω τον Αναπληρωτή Καθηγητή Συμεών Παπαβασιλείου που, ως μέλος της Τριμελούς Επιτροπής μου και ως Αναπληρωτής Διευθυντής του NETMODE, ήταν εκεί για την παροχή κάθε είδους συμβουλής, με την γνώμη του να είναι σημείο αναφοράς.

Η συνεισφορά έχει καμιά φορά και την όψη των γυμνασίων, ακούσιων και εκούσιων, ή την όψη του Ερευνητή Β' Ε.Π.Ι.Σ.Ε.Υ Δημήτρη Καλογερά. Το 3^{ο} μέλος της Τριμελούς Επιτροπής, σε προπονητικό ρόλο, είχε αναλάβει κατά την διάρκεια της εκπόνησης της Διατριβής να σηκώνει τον κυματισμό στον οποίο αναφέρθηκα

πρωτύτερα. Κάπου ανάμεσα στην ωκεανοπλοΐα του ερευνητικού πεδίου και στους σωματικούς αγώνες αντοχής με έμαθε να στέκομαι, στα πόδια μου μετά από 42 χιλιόμετρα δρόμου. Με έμαθε ότι οι αγώνες αντοχής είναι οι μόνοι που μοιάζουν με την ίδια την ζωή, θέλουν στρατηγικό χειρισμό των αποθεμάτων πάσης φύσεως και το μυαλό να λειτουργεί όταν πλέον το σώμα και η ψυχή δεν υπακούουν. Αποτέλεσε πηγή έμπνευσης για την εκπλήρωση του στόχου του Διδακτορικού, αλλά και του Κλασσικού Μαραθωνίου, μαζί με τους υπόλοιπους δρομείς μεγάλων αποστάσεων του NETMODE (Τρύφωνα Χιώτη, Γιώργο Κουτέπα, Βασίλη Μερεκούλια).

Η διατήρηση της ερευνητικής πορείας, εκτός από φάρους, απαιτεί και πλήρωμα στο πλεούμενο. Όλοι τους στο NETMODE ήταν εξαιρετικής ποιότητας επιστήμονες και ηθικά στοιχεία, που τους αξίζει ένα μεγάλο ευχαριστώ για την σύμπλευση: Μαίρη Γραμματικού, Λεωνίδας Λυμπερόπουλος , Βαγγέλης Ανυφαντής, Βασίλης Μερεκούλιας, Γιώργος Κατσίνης, Χρύσα Παπαγιάννη, Βασίλης Καρυώτης, Τιμόθεος Καστρινογιάννης, Κώστας Τρούλος, Άγγελος Λένης, Γιώργος Αριστομενόπουλος, Στέλλα Καφετζόγλου, Κώστας Μαρίνος, Βασιλική Πουλή, Αλέξανδρος Σιούγγαρης, Έλενα Στάη, Άρης Λειβαδέας, Γιάννος Κρύφτης, Άγγελος Καπουκάκης, Μαίρη Γιατίλη. Αλλά και παλαιότερους συντελεστές του NETMODE, που τους συνάντησα σε ρόλους πλέον διαφορετικούς, καταξιωμένους και έμπειρους, όπως ο Γιώργος Κουτέπας, ο Χρήστος Στάθης, ο Τρύφων Χιώτης, ο Φώτης Σταματελόπουλος αποτέλεσαν φωτεινά παραδείγματα των «πρώτων» του NETMODE.

Δεν μπορώ να ξεχάσω τους συναδέλφους που έγιναν φίλοι και συγχρόνως μέντορες στον κόσμο της επαγγελματικής μου κατάρτισης και ακούν στο όνομα Παναγιώτης Χριστιάς και Θανάσης Δουΐτσης. Ακρογωνιαίοι λίθοι στο Κέντρο Διαχείρισης Δικτύων του Ε.Μ.Π, κάτοχοι αξιοζήλευτης τεχνικής αρτιότητας και ήθους υψηλότερου από το μπόι του Παναγιώτη, που δίχως να ακούγονται, προσφέρουν τα μέγιστα, σε όλα τα επίπεδα.

Στην συνέχεια, θα ήθελα να αναφερθώ στους συνεργάτες, που όχι απλά συμπορεύτηκα μαζί τους, αλλά με έκαναν να μην χάνω τον βηματισμό, ξεκινώντας με τον άνθρωπο που διαδραμάτισε τον πιο καταλυτικό ρόλο στην ακαδημαϊκή μου εξέλιξη, τον Γιώργο Ανδρουλιδάκη. Ως μεταδιδακτορικός ερευνητής υπήρξε αυτός που μου έδειξε πως γίνεται η σύνθεση των μικρών βημάτων σε μεγαλύτερα και πώς να ορίζω σημεία καμπής σε μια πορεία που μοιάζει ατελείωτη. Δίπλα μου όμως στάθηκαν και

μικρότεροι σε ηλικία, φοιτητές της Σχολής Ηλεκτρολόγων και Μηχανικών Υπολογιστών του Ε.Μ.Π, που μοιράσθηκαν μαζί μου τις αγωνίες και την προσπάθεια: Νίκος Λουτζάκης, Μάνος Δημογεροντάκης, Κώστας Γιώτης, Σπύρος Μαστοράκης. Οι αμέτρητες ώρες κοινής δουλειάς, οι εντατικοί ρυθμοί που ακολουθήσαμε και οι κοινές θυσίες νομίζω πως δικαιώθηκαν από την πρόοδο που επέδειξαν.

Για το τέλος άφησα αυτούς που εξασφάλισαν την ερμάτωσή μου σε θάλασσες δύσκολες. Είναι η οικογένεια μου, ο πατέρας μου, Γιώργος, η μητέρα μου, Τόνια και ο αδελφός μου, Μιχάλης, που μαζί με τους φίλους μου Κωνσταντίνο, Γεωργία, Γιώργο, Θανάση, Λάμπη, Χρήστο ήταν κοντά μου, να εξασφαλίζουν την απαιτούμενη ευστάθεια και να βοηθούν, ο καθένας με τον τρόπο που γνώριζε.

Πίνακας Περιεχομένων

Ευρετήριο Σχημάτων

Ευρετήριο Πινάκων

1 Εισαγωγή

1.1 Γενικό Πλαίσιο

Το οικοσύστημα του Διαδικτύου του Μέλλοντος (Future Internet – FI) αναπτύσσεται με γνώμονα την ευελιξία σε επίπεδο αρχιτεκτονικής αποσκοπώντας στην υποστήριξη νέων πρωτοκόλλων επικοινωνίας και τεχνολογιών. Πρωταρχικός στόχος είναι μια εκ βάθρων αναθεώρηση του δικτυακού κόσμου, όπως είχε οριστεί μέχρι σήμερα. Οι τεχνολογίες εικονικοποίησης (network virtualization), νέα πρωτόκολλα ελέγχου των δικτύων, όπως το OpenFlow, και οι νέες σχεδιαστικές αρχές που αναπτύσσονται γύρω από τις συγκεκριμένες τεχνολογίες αποτελούν κομμάτι του οικοσυστήματος του Διαδικτύου του μέλλοντος.

Οι ερευνητικές προσπάθειες εξέλιξης και επανασχεδιασμού εκ βάθρων (clean-slate approach) των δικτυακών πρωτοκόλλων και τεχνολογιών καθώς και οι τεχνολογίες εικονικοποίησης πόρων οδηγούν στην δημιουργία νέων πειραματικών δικτυακών υποδομών για το Διαδίκτυο του μέλλοντος (Future Internet). Τα βασικά χαρακτηριστικά της νέας γενιάς πειραματικών υποδομών είναι (α) η δυνατότητα χρήσης τους από πολλαπλούς ερευνητές, (β) οι αυξημένες δυνατότητες ελέγχου τους από τον ερευνητή, (γ) η δυναμική παροχή πόρων στους ερευνητές, (δ) η προσπάθεια διασύνδεσης υποδομών για την εκτέλεση πειραμάτων μεγάλης κλίμακας.

Οι πρώτες υποδομές πειραματισμού για το Διαδίκτυο του μέλλοντος (Future Internet Infrastructures - FI) περιλαμβάνουν το PlanetLab [Chun03], το Virtual Network Infrastructure [Bavie06] , το EmuLab [Hible08], το ProtoGeni [ProtoGeni] υπό την ερευνητική ομπρέλα του Global Environment for Network Innovations (GENI) [GENI] στις Ηνωμένες Πολιτείες. Στην Ευρώπη αντίστοιχα περιλαμβάνουν το PanLab [PanLab], το FEDERICA [FEDERICA], το OneLab [OneLab] το OFELIA [OFELIA] καθώς και μια πληθώρα άλλων ερευνητικών προσπαθειών υπό την ομπρέλα του Future Internet and Research Experimentation (FIRE) [FIRE].

Η προσπάθεια δημιουργίας ομοσπονδιών πειραματικών δικτυακών υποδομών ικανών να υποστηρίζουν πολλαπλά πειράματα, με χρήση τεχνολογιών εικονικοποίησης πόρων, εστιάζεται κατά κύριο λόγο στην ανάπτυξη εργαλείων και διεπαφών

προγραμμάτων (program interfaces) που θα επιτρέπουν την ομοσπονδοποίηση υποδομών. Η πρώτη προσπάθεια προς αυτή την κατεύθυνση ήταν η ανάπτυξη της αρχιτεκτονικής Slice-based Federation Architecture (SFA), που υιοθετήθηκε από το GENI στις Ηνωμένες Πολιτείες και από το OneLab στην Ευρώπη καθώς και η ανάπτυξη της αρχιτεκτονικής Teagle [Teagle] που χρησιμοποιήθηκε από το PanLab.

Η ομοσπονδοποίηση των υποδομών για τους ερευνητές του σήμερα και τους χρήστες του μέλλοντος σημαίνει ότι θα έχουν την δυνατότητα χρήσης πόρων από πολλαπλές υποδομές αυξάνοντας το μέγεθος και την ποικιλότητα των πειραματικών διατάξεων τους. Η χρήση τεχνολογιών εικονικοποίησης σε επίπεδο κόμβων (compute virtualization) αποτελεί βασική προϋπόθεση για την δυναμική παροχή πόρων και την ταυτόχρονη χρήση τους.

1.2 Διατύπωση του Προβλήματος

Η δημιουργία υποδομών εικονικών πόρων έχει διευρύνει το πλαίσιο δυνατοτήτων και ευελιξίας του χρήστη, ανοίγοντας τον δρόμο για αυξημένο έλεγχο πάνω στους αποδιδόμενους πόρους και στα τρία επίπεδα δικτύωσης: (α) επίπεδο προώθησης δεδομένων, (β) επίπεδο διαχείρισης, (γ) επίπεδο ελέγχου. Η εργασία μας πραγματεύεται επιμέρους ανοιχτά προβλήματα και στα τρία επίπεδα:

Επίπεδο Προώθησης Δεδομένων

Ανοιχτό ερευνητικό θέμα αποτελεί η δημιουργία μηχανισμών δυναμικής απόδοσης πόρων, ελεγχόμενων από το χρήστη, για την δημιουργία ζεύξεων/τοπολογιών σε ετερογενείς ομόσπονδες πειραματικές υποδομές εικονικών πόρων.

Στόχος μας είναι η δικτυακή διασύνδεση εικονικών υπολογιστικών οντοτήτων (virtual nodes) οι οποίες βρίσκονται σε ομόσπονδες διαχειριστικές περιοχές (federated domains) ετερογενών υποδομών.

Η ερευνητική μας εργασία εστιάζει στην μελέτη και δημιουργία ενός μηχανισμού διασύνδεσης ετερογενών ομόσπονδων υποδομών εικονικοποιημένων πόρων για την ταυτόχρονη εξυπηρέτηση πολλαπλών χρηστών.

Το θέμα της δυναμικής απόδοσης δικτυακών και υπολογιστικών πόρων σε ομοσπονδίες ετερογενών πειραματικών υποδομών αποτέλεσε αντικείμενο έρευνας του ερευνητικού προγράμματος Network Innovation over Virtualized Infrastructures (NOVI) [NOVI] [Lymp12] στο οποίο εφαρμόσαμε την προτεινόμενη λύση.

Επίπεδο Διαχείρισης Δικτύου

Η παρακολούθηση και η συλλογή στατιστικών της δικτυακής κίνησης στις υποδομές που αναπτύχθηκαν [ProtoGeni] [GEANTOF], αρχικά, αφέθηκε στις εγγενείς δυνατότητες του πρωτοκόλλου OpenFlow. Η συγκεκριμένη τακτική μεταφέρει την αρμοδιότητα παρακολούθησης της δικτυακής κίνησης από το επίπεδο διαχείρισης στο επίπεδο ελέγχου.

Ερευνούμε κατά πόσο ένα πρωτόκολλο κεντρικοποιημένου ελέγχου όπως το OpenFlow μπορεί ταυτόχρονα να μεταφέρει όλες τις απαιτούμενες εντολές προώθησης ροών στις δικτυακές συσκευές και συγχρόνως να ικανοποιεί τις ανάγκες παθητικής παρακολούθησης, όπως αρχικά είχε υποστηριχθεί από την ερευνητική κοινότητα [Brag10] [Mehd11].

Ελέγχουμε σε πειραματικό επίπεδο την αποτελεσματικότητα και την κλιμακοθετησιμότητα του OpenFlow στη λειτουργία της παθητικής παρακολούθησης του δικτύου για απαιτητικές εφαρμογές, όπως είναι ο εντοπισμός δικτυακών ανωμαλιών (anomaly detection). Επίσης ελέγχουμε την ικανότητα αντιμετώπισης τέτοιων ανωμαλιών με χρήση του OpenFlow.

Προδιαγράφουμε πλαίσιο παθητικής παρακολούθησης υποδομών δικτυακής εικονικοποίησης ελεγχόμενες με το πρωτόκολλο OpenFlow που κάνει χρήση του sFlow και βοηθητικά άλλων πρωτοκόλλων όπως το SNMP.

Επίπεδο Ελέγχου

Για την απόδοση μεγαλύτερου ελέγχου των δικτυακών πόρων (εικονικοποίηση του επιπέδου ελέγχου των δικτύων), πολλές ερευνητικές προσπάθειες στηρίχθηκαν πάνω σε ένα νέο πρωτόκολλο του επιπέδου ελέγχου που ονομάζεται OpenFlow [McKe08]. Υποδομές όπως το ProtoGeni, το OFELIA, το GÉANT OpenFlow Facility [GEANTOF] επέτρεψαν την απόδοση μέρους του επιπέδου ελέγχου των πειραματικών δικτύων στον χρήστη, στηριζόμενες σε τεχνολογίες εικονικοποίησης όπως ο FlowVisor [Sher10] και το OpenFlow.

Το ζήτημα της δυναμικής απόδοσης λειτουργιών του επιπέδου ελέγχου σε πολλαπλούς χρήστες εικονικών δικτύων διαμοιραζόμενων υποδομών παραμένει ανοιχτό. Σε τέτοιες υποδομές απαιτείται ταυτόχρονη απόδοση του ελέγχου εύρους δικτυακών ροών (flowspace) σε διαφορετικούς χρήστες, με διατήρηση της απομόνωσης. Ο χρήστης θα πρέπει να έχει την ικανότητα να επιβάλλει την δική του λογική δικτυακής διασύνδεσης, οριζόμενης από λογισμικό (SDN) μόνο στους πόρους που του έχουν αποδοθεί.

Η ανάθεση μέρους του flowspace σε χρήστες απαιτεί την ικανότητα ελέγχου πιθανών συγκρούσεων μεταξύ αιτημάτων διαφορετικών χρηστών (flowspace conflict detection) καθώς και την ανάλυση των αιτημάτων ως προς την συμβατότητα τους με την λογική διαχωρισμού (slicing) που θέλει να επιβάλλει ο διαχειριστής της υποδομής (flow analysis) [FOAM]. Η ερευνητική μας εργασία έχει ως σκοπό να προσδιορίσει και να αξιολογήσει πολιτικές διαμοιρασμού του flowspace που εξασφαλίζουν την απομόνωση των πόρων που διαχειρίζονται οι χρήστες και αποτελούν τα εικονικά τους δίκτυα.

1.3 Παρουσίαση Περιεχομένων

Στο **Κεφάλαιο 2** γίνεται παρουσίαση των επιπέδων λειτουργιών σε μοντέλα αναφοράς αρχιτεκτονικής δικτύου και δίνεται βάρος στο κεντρικοποιημένο επίπεδο ελέγχου, όπως αυτό έχει διαμορφωθεί με την χρήση του πρωτοκόλλου OpenFlow σε δίκτυα οριζόμενα από λογισμικό (SDN). Στο **Κεφάλαιο 3** ακολουθεί αναφορά στην εικονικοποίηση δικτύων υπολογιστών και μελετώνται οι σχεδιαστικές ελλείψεις των πρώιμων υλοποιήσεων εικονικοποίησης καθώς και η εξέλιξή τους.

Η σχεδίαση και τα πειραματικά αποτελέσματα του πρωτότυπου μηχανισμού διασύνδεσης (stitching) πολλαπλών εικονικών υπολογιστικών πόρων ετερογενών ομόσπονδων υποδομών στο επίπεδο προώθησης δεδομένων περιγράφονται στο **Κεφάλαιο 4 .** Γίνεται ειδική αναφορά στη δυναμική απόδοση πόρων σε ομοσπονδίες ετερογενών πειραματικών υποδομών, όπως αυτή του ερευνητικού προγράμματος Network Innovation over Virtualized Infrastructures (NOVI) στο οποίο εφαρμόσαμε την προτεινόμενη λύση.

Εν συνεχεία στο **Κεφάλαιο 5**, μελετάται η μέθοδος παθητικής παρακολούθησης με χρήση των εγγενών δυνατοτήτων του πρωτοκόλλου OpenFlow και των μειονεκτημάτων που παρουσιάζει η παρακολούθηση της δικτυακής κίνησης με την χρήση ενός πρωτοκόλλου κεντρικοποιημένου ελέγχου. Ακολουθεί η περιγραφή πλαισίου παθητικής παρακολούθησης υποδομών δικτυακής εικονικοποίησης που κάνει χρήση των πρωτοκόλλων OpenFlow και sFlow και έχει ως στόχο την ικανότητα χειρισμού μεγαλύτερου αριθμού ροών και ρυθμών μετάδοσης δεδομένων, σε σχέση με λύσεις που βασίζονται αποκλειστικά στο OpenFlow. Παρουσιάζονται πειραματικές μετρήσεις που προσφέρει ο μηχανισμός παθητικής παρακολούθησης δικτύων οριζόμενων από λογισμικό με και χωρίς την συνδρομή του sFlow στον εντοπισμό και εξομάλυνση των δικτυακών ανωμαλιών. Βαρύτητα δίνεται στην αποτελεσματικότητα του sFlow (ευστοχία εντοπισμού δικτυακών ανωμαλιών, κατανάλωση πόρων του επιπέδου ελέγχου) και στην κλιμακοθετησιμότητα.

Στο **Κεφάλαιο 6** παρουσιάζεται μηχανισμός ανάθεσης ροών ανά εικονικό δίκτυο χρήστη σε περιβάλλοντα SDN κεντρικοποιημένου ελέγχου που κάνουν χρήση του πρωτοκόλλου OpenFlow. Ιδιαίτερη μελέτη, μέσω της υλοποίησης πρωτότυπης μηχανής διαμοιρασμού του πεδίου ορισμού των ροών (flowspace) και πειραματικών

μετρήσεων, γίνεται για να διαπιστωθεί πως περιβάλλοντα SDN κεντρικοποιημένου ελέγχου μεγάλης κλίμακας μπορούν να κάνουν χρήση του πρωτοκόλλου OpenFlow για την παροχή προηγμένων υπηρεσιών εικονικής δικτύωσης σε πολλαπλούς χρήστες.

Το **Κεφάλαιο 7** έχει αφιερωθεί για τη σύνοψη των συμπερασμάτων καθώς και την παρουσίαση επιστημονικών προεκτάσεων που αξιολογήθηκαν ως υψηλής ερευνητικής αξίας κατά τη συγγραφή της παρούσας Διδακτορικής Διατριβής. Τέλος, στο **Κεφάλαιο 8** συγκεντρώνονται οι αναφορές που σχετίζονται με την παρούσα Διδακτορική Διατριβή και στο **Κεφάλαιο 9** αναφέρονται, συνοπτικά, τα στοιχεία των δημοσιεύσεων σε Περιοδικά και Συνέδρια με κρίση που προέκυψαν στα πλαίσια της παρούσας εργασίας.

2 Επίπεδα Λειτουργιών σε Μοντέλα Αναφοράς Αρχιτεκτονικής Δικτύου

Functional Planes in Network Architecture Reference Models

Στις κλασσικές προτυποποιημένες αρχιτεκτονικές δικτύωσης, η κάθε δικτυακή συσκευή αποτελούσε μια αυτόνομη οντότητα. Εσωτερικά της, για λόγους οργάνωσης και απόδοσης, οι απαιτούμενοι μηχανισμοί ήταν σχεδιασμένοι και υλοποιημένοι σε διακριτές ομάδες με βασικότερες αυτές που αφορούσαν τις **λειτουργίες προώθησης των δεδομένων (forwarding/data plane)**, τις **λειτουργίες διαχείρισης (management)** καθώς και τις **λειτουργίες ελέγχου (control)** [ITU-I322] [Rexf04].

Οι μηχανισμοί που υλοποιούσαν τις προαναφερθείσες ομάδες λειτουργιών χρησιμοποιούσαν ιδιοταγείς (proprietary) διεπαφές μεταξύ τους που αναπτύσσονταν από κάθε κατασκευαστή δικτυακών συσκευών χωριστά, χωρίς καμία δυνατότητα επιλογών, ειδικά στην περίπτωση της διεπαφής μεταξύ λειτουργιών ελέγχου και προώθησης δεδομένων. Το συγκεκριμένο γεγονός ήρθε να ανατρέψει η εμφάνιση αρχιτεκτονικών όπως είναι η ForCES [RFC3746] και το πρωτόκολλο OpenFlow, το οποίο αποτελεί και αντικείμενο μελέτης της διατριβής και αναλύεται στα βασικά του σημεία στην Παράγραφο 2.4.

2.1 Επίπεδο Προώθησης Δεδομένων (Forwarding/Data-plane)

Το επίπεδο προώθησης δεδομένων, παραδοσιακά, υλοποιείται τοπικά σε κάθε δικτυακή συσκευή και λειτουργεί με την ταχύτητα άφιξης των πακέτων (line-rate).

Μια βασική λειτουργία, που υλοποιείται σε δρομολογητές για παράδειγμα, είναι η προώθηση των πακέτων σε κάποια διεπαφή εξόδου (egress interface) βάσει κάποιων κανόνων δρομολόγησης (π.χ longest-prefix match) και ο ταυτόχρονος έλεγχος πλήρωσης κριτηρίων φιλτραρίσματος (Access Control Lists - ACLs).

Στο συγκεκριμένο επίπεδο υλοποιούνται επίσης λειτουργίες ελέγχου ροής πακέτων με χρήση ουρών (queue management) και προγραμματισμού πακέτων (packet

scheduling). Το κύριο χαρακτηριστικό είναι ότι οι προαναφερθείσες λειτουργίες υλοποιούνται με χρήση εξειδικευμένου υλισμικού (hardware).

Αν και οι υλοποιήσεις του επιπέδου προώθησης πακέτου διαφέρουν ανά κατασκευαστή, η επικοινωνία μεταξύ συσκευών διαφορετικών κατασκευαστών είναι δυνατή λόγω των τυποποιημένων (standardized) πρωτοκόλλων προώθησης δεδομένων (π.χ Ethernet, Internet Protocol).

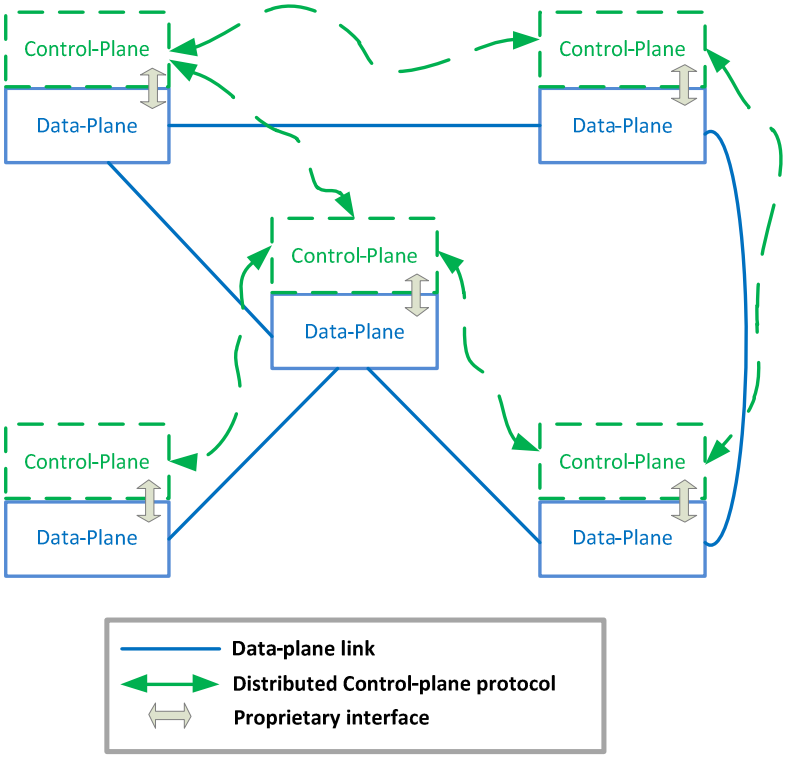
2.2 Επίπεδο Ελέγχου (Control-Plane)

Το επίπεδο ελέγχου είναι υπεύθυνο να καθορίζει τη συμπεριφορά του επιπέδου προώθησης δεδομένων εφαρμόζοντας κατάλληλους αλγορίθμους. Η λειτουργία του επιπέδου μπορεί να είναι κεντρικοποιημένη (centralized) ή κατανεμημένη (distributed). Στην περίπτωση που είναι κεντρικοποιημένη, οι αποφάσεις λαμβάνονται σε ένα σημείο, αλλά αφορούν όλο το δίκτυο, ενώ στην περίπτωση της κατανεμημένης λειτουργίας του επιπέδου ελέγχου οι αλγόριθμοι που χρησιμοποιούνται εφαρμόζονται σε κάθε δικτυακή συσκευή που λαμβάνει μέρος στο επίπεδο ελέγχου.

Μια από τις κύριες λειτουργίες του επιπέδου ελέγχου είναι να υπολογίζει διαδρομές δεδομένων συνδυάζοντας πληροφορίες από κάθε πρωτόκολλο δρομολόγησης, δημιουργώντας το Forwarding Information Base (FIB) που τελικά χρησιμοποιείται από το επίπεδο προώθησης δεδομένων.

Κατανεμημένο Επίπεδο Ελέγχου (Distributed Control-Plane)

Οι αλγόριθμοι κατανεμημένου ελέγχου απαιτούν την ανταλλαγή πληροφοριών μεταξύ των δικτυακών συσκευών που συμμετέχουν στον έλεγχο της προώθησης πακέτων. Τα πρωτόκολλα του κατανεμημένου επιπέδου ελέγχου περιγράφουν τη διαδικασία ανταλλαγής μηνυμάτων καθώς και τον αλγόριθμο που χρησιμοποιείται για την λήψη των αποφάσεων (**Σχήμα 1**).



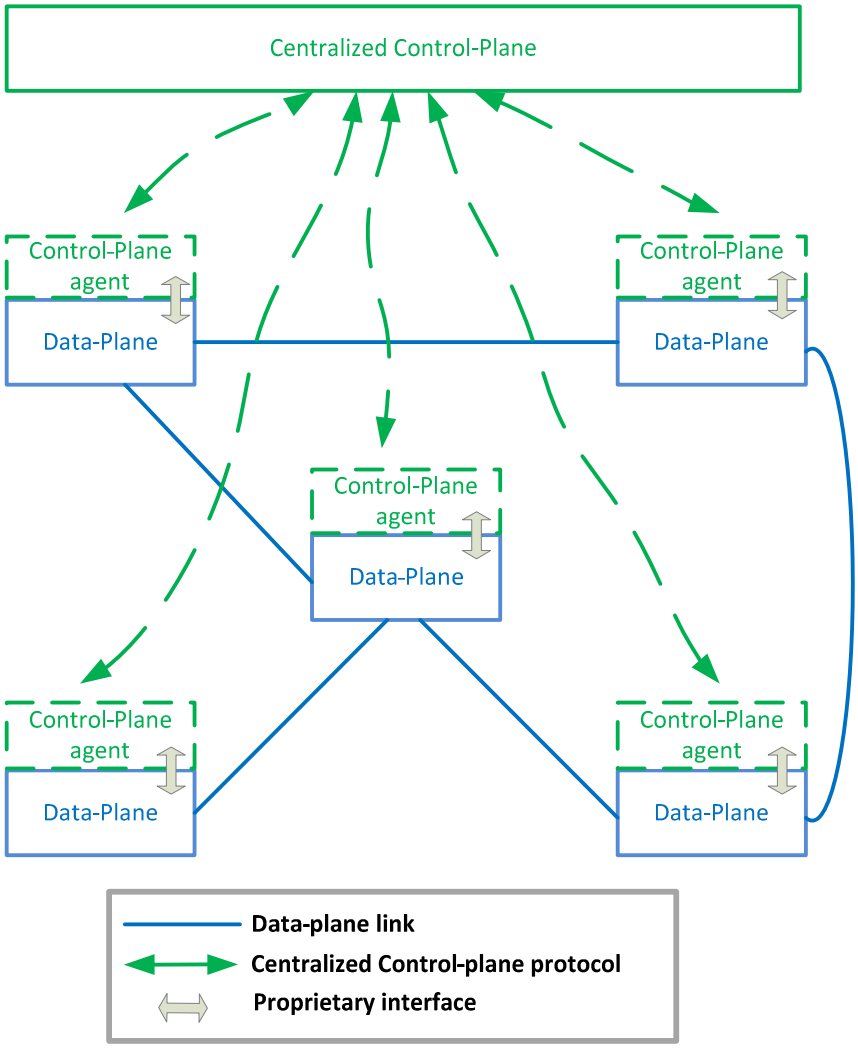
Σχήμα 1: Κατανεμημένο επίπεδο ελέγχου

Για παράδειγμα τα μηνύματα ενημέρωσης (update messages) και η διαδικασία λήψης αποφάσεων δρομολόγησης του πρωτοκόλλου διατομεακής δρομολόγησης (inter-domain routing) Border Gateway Protocol (BGP) είναι χαρακτηριστικά στοιχεία του κατανεμημένου επιπέδου ελέγχου. Ενδεικτικά αναφέρουμε ότι στο συγκεκριμένο επίπεδο ανήκουν και οι ενημερώσεις για την κατάσταση διεπαφών (Link State Advertisements - LSAs) και ο αλγόριθμος Dijkstra που αποτελούν μέρος του πρωτοκόλλου Open Shortest Path First (OSPF).

Στην συντριπτική πλειοψηφία των δικτυακών συσκευών (π.χ. δρομολογητές) που το επίπεδο ελέγχου είναι κατανεμημένο, η επικοινωνία μεταξύ του επιπέδου ελέγχου της εκάστοτε συσκευής με το επίπεδο προώθησης δεδομένων γίνεται ακόμα και σήμερα μέσω ιδιοταγών (proprietary) διεπαφών. Η συγκεκριμένη χρήση ιδιοταγών διεπαφών καθιστά τις υλοποιήσεις των δικτυακών συσκευών ένα κλειστό σιλό τεχνολογιών ενός κατασκευαστή, μια προσέγγιση, που στον κόσμο των υπολογιστικών συστημάτων γενικής χρήσης έχει ξεπεραστεί εδώ και δυο δεκαετίες.

Κεντρικοποιημένο Επίπεδο Ελέγχου (Centralized Control-Plane)

Την τελευταία δεκαετία, άρχισαν να κερδίζουν έδαφος προσεγγίσεις που προωθούσαν την αποσύζευξη (decoupling) του επιπέδου ελέγχου με αυτό της προώθησης δεδομένων (**Σχήμα 2**). Στην κυριαρχούσα κεντρικοποιημένη αρχιτεκτονική επιπέδου ελέγχου υπάρχει μια κεντρική οντότητα που υπαγορεύει τον τρόπο προώθησης των πακέτων στο δίκτυο. Το δίκτυο προγραμματίζεται με την χρήση μιας ομάδας εντολών που έχει ως σκοπό την εισαγωγή ροών που χρησιμοποιούνται από το επίπεδο προώθησης πακέτων των συσκευών. Με αυτό τον τρόπο η διαχείριση και η λειτουργία των δικτυακών συσκευών γίνεται απλούστερη, σε αντιδιαστολή με αυτές των κεντρικών οντοτήτων ελέγχου που συγκεντρώνουν την πολυπλοκότητα.



Σχήμα 2: Κεντρικοποιημένο επίπεδο ελέγχου

Παράλληλα με το διαχωρισμό λειτουργιών ελέγχου και προώθησης δεδομένων άρχισε να προωθείται και ο ορισμός ανοιχτών/προτυποποιημένων (open/standardized) διεπαφών, κατά αντιστοιχία με αυτές που είχαν ήδη δημιουργηθεί στα λειτουργικά συστήματα υπολογιστών [Hjal00] [Pete00] [Kohl00]. Το πρώτο γενικευμένο πλαίσιο που αναπτύχθηκε ήταν το Forwarding and Control Element Separation (ForCES) Framework, που όριζε την διαίρεση των δικτυακών οντοτήτων (network elements) σε οντότητες ελέγχου (control elements) και προώθησης δεδομένων (forwarding elements) καθώς και τις αναμεταξύ τους διεπαφές.

2.3 Επίπεδο Διαχείρισης (Management-Plane)

Για το επίπεδο διαχείρισης δικτύου έχουν αναπτυχθεί πολλαπλά μοντέλα τα 30 τελευταία χρόνια από διαφορετικά σώματα τυποποίησης όπως είναι ο Διεθνής Οργανισμός Τυποποίησης (International Organization for Standardization - ISO), η Διεθνής Ένωση Τηλεπικοινωνιών (International Telecommunication Union - ITU) καθώς και η Ομάδα Έργου Μηχανίκευσης Ίντερνετ (Internet Engineering Task Force – IETF).

Το δημοφιλέστερο μοντέλο είναι το FCAPS, που αρχικά ορίστηκε από τον ISO [OSI] και μετέπειτα υιοθετήθηκε και τροποποιήθηκε σε ορισμένα σημεία από την ITU [M3400].

Συνοπτικά αναφέρουμε τις 5 ομάδες διαχείρισης που ορίζονται στο επίπεδο διαχείρισης στο μοντέλο FCAPS. Οι 5 ομάδες διαχείρισης χρησιμοποιήθηκαν και από την IETF [RFC5951] για τα Multi-Protocol Label Switching (MPLS) δίκτυα.

Τα πέντε επίπεδα διαχείρισης παρουσιάζονται συνοπτικά παρακάτω:

- Διαχείριση Βλαβών/Λαθών – Fault Management

  Ο εντοπισμός βλαβών, η γνωστοποίησή τους, η απομόνωσή τους για την μείωση ή την εξάλειψη της επίδρασής τους στην ορθή λειτουργία των δικτυακών υπηρεσιών καθώς και η διόρθωσή τους.

- Διαχείριση Διάρθρωσης – Configuration Management

Η διάρθρωση των παραμέτρων των δικτυακών συσκευών αποτελεί ένα από τα σημαντικότερα κομμάτια διαχείρισης στα δικτυακά περιβάλλοντα που λαμβάνει μέρος κατά την αρχική εγκατάσταση των συσκευών αλλά και κατά την διάρκεια της λειτουργίας τους ανάλογα με τις εκάστοτε ανάγκες που παρουσιάζονται για την ορθή λειτουργία των υπηρεσιών καθώς και για την παροχή τους.

- Λογιστική Διαχείριση – Accounting Management

Η συλλογή και η αποθήκευση δεδομένων χρήσης των πόρων (resource usage), με σκοπό την κοστολόγηση των υπηρεσιών και την λογιστική διαχείρισής τους.

- Διαχείριση επιδόσεων – Performance Management

Η συλλογή και η αποθήκευση των στατιστικών λειτουργίας όσον αφορά τη χρήση των πόρων, με σκοπό την βελτιστοποίηση της λειτουργίας και τον σχεδιασμό/πρόβλεψη μελλοντικών αναγκών.

- Διαχείριση ασφαλείας – Security Management

Ο έλεγχος της πρόσβασης στις παρεχόμενες υπηρεσίες και οι μηχανισμοί ασφαλείας που σχετίζονται με την πρόσβαση σε λειτουργίες του επιπέδου διαχείρισης και ελέγχου των δικτυακών πόρων.

2.4 Πρωτόκολλο OpenFlow (OpenFlow Protocol)

Το Forwarding and Control Element Separation (ForCES) Framework είχε εισάγει την ιδέα του διαχωρισμού μεταξύ επιπέδου ελέγχου και επιπέδου προώθησης πακέτων. Στη συνέχεια το OpenFlow [McKe08] καθόρισε μια ολοκληρωμένη διεπαφή μεταξύ των δυο επιπέδων [OFspec10]. Κάνοντας το συγκεκριμένο διαχωρισμό επέτρεψε την ταχύτερη ανάπτυξη των οριζόμενων από λογισμικό δικτύων (Software Defined Networking – SDN) [ONF12], που αποτελεί ένα από τα κύρια αντικείμενα ενασχόλησης του Open Networking Foundation (ONF).

Η αρχική ανάπτυξη του πρωτοκόλλου ξεκίνησε στον ακαδημαϊκό χώρο. Το ONF μετέπειτα ανέλαβε την εξέλιξη του πρωτοκόλλου και ολόκληρου του οικοσυστήματος

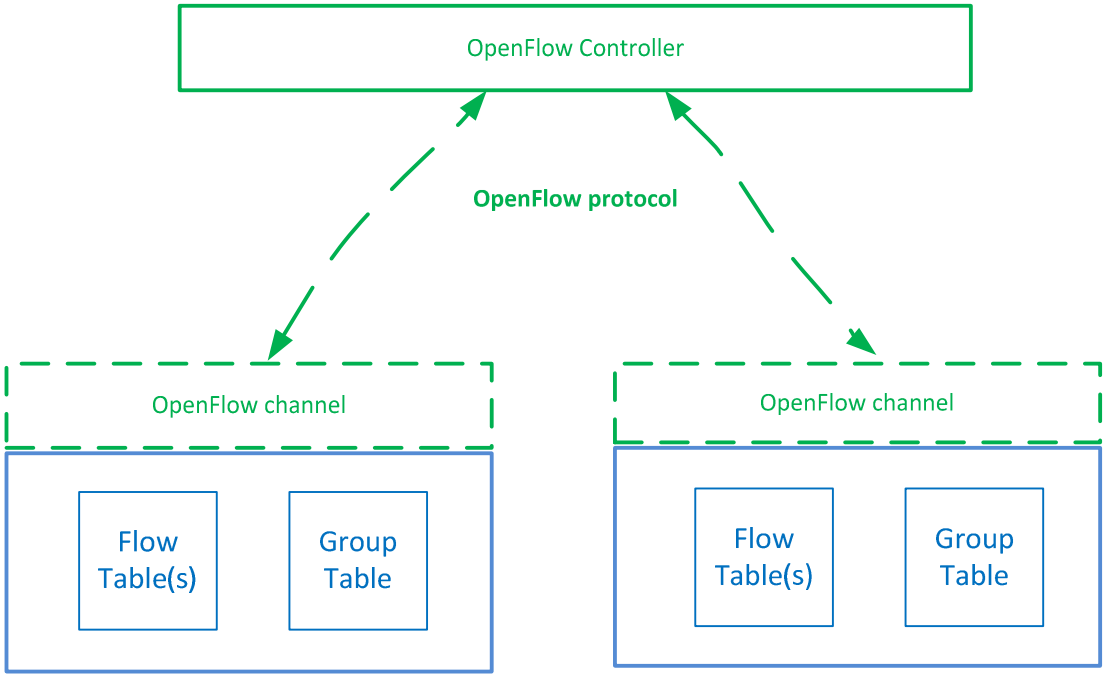
δικτύωσης που αναπτύσσεται γύρω του, όπως οι δοκιμές και οι πιστοποιήσεις υλισμικού για τη συμβατότητά τους με το OpenFlow specification.

Το OpenFlow μπορεί να θεωρηθεί το ανάλογο του συνόλου εντολών (instruction set) ενός επεξεργαστή για δικτυακές συσκευές. Δίνει την δυνατότητα σε λογισμικό να προγραμματίσει τους δικτυακούς επεξεργαστές που συμμετέχουν στο επίπεδο προώθησης πακέτων, ανεξαρτήτως της εσωτερικής αρχιτεκτονικής των δικτυακών επεξεργαστών.

Ένα βασικό χαρακτηριστικό του είναι ότι όσες συσκευές είναι συμβατές μπορούν να υλοποιούν προώθηση πακέτων λαμβάνοντας υπόψη πολλαπλές επικεφαλίδες πρωτοκόλλων διαφορετικών επιπέδων της στοίβας πρωτοκόλλων (network protocol stack). Η συγκεκριμένη ιδιότητα επιτρέπει την προώθηση ροών (flow forwarding) με σύνθετα κριτήρια και όχι απλώς την προώθηση πακέτων (packet forwarding) βάσει των επικεφαλίδων ενός συγκεκριμένου πρωτοκόλλου (π.χ Internet Protocol). Η συγκεκριμένη δυνατότητα δινόταν μέχρι σήμερα από διάφορους κατασκευαστές, αλλά δεν υπήρχε κάποια ευρέως υιοθετημένη λύση.

Τεχνική Περιγραφή OpenFlow (OpenFlow Specification)

Μια δικτυακή συσκευή του επιπέδου προώθησης δεδομένων που υποστηρίζει το OpenFlow protocol ονομάζεται **OpenFlow switch**, σύμφωνα με το OpenFlow Specification. Τα OpenFlow switches ελέγχονται μέσω του δικτυακού πρωτοκόλλου **OpenFlow protocol** από μια ανεξάρτητη συσκευή που ανήκει στο επίπεδο ελέγχου και ονομάζεται **OpenFlow Controller**, όπως απεικονίζεται στο (**Σχήμα 3**). Ο OpenFlow controller μπορεί να ελέγχει τον μηχανισμό προώθησης πακέτων που βρίσκεται υλοποιημένος μέσα στα OpenFlow switches καθορίζοντας τους κανόνες των πινάκων προώθησης που ονομάζονται **Flow tables**. Σε περίπτωση που τα Flow tables είναι πολλαπλά, η προσπέλασή τους είναι μονόδρομη. Η προσπέλασή τους ,δηλαδή, κατά την διαδικασία ταύτισης (matching) ενός πακέτου, γίνεται μόνο προς τα εμπρός, δίχως δυνατότητα μεταπήδησης από επόμενο flow table σε προηγούμενο.



Σχήμα 3: Κεντρικοποιημένο επίπεδο ελέγχου βασιζόμενο στο OpenFlow

Τα OpenFlow switches θα πρέπει να έχουν την ικανότητα να προωθούν Ethernet frames βάσει ενός συνόλου κανόνων που είναι αποθηκευμένοι σε ένα ή περισσότερα Flow tables. Τα βασικά στοιχεία κάθε εγγραφής (**Flow entry**) στον Flow table είναι :

- ένα πεδίο ταύτισης (**Match Field)**
- ένα σύνολο κανόνων (**Actions)** που καθορίζουν την τύχη του πακέτου
- μετρητές (**Counters**) που ανανεώνονται κάθε φορά που ένα πακέτο ταιριάζει στο πεδίο ταύτισης
- ένα πεδίο προτεραιότητας του flow entry (**Priority**) και ένα πεδίο χρονικού ορίου λήξεως της ισχύος του κανόνα (**Time-out**).

Τα πεδία ταύτισης, οι κανόνες και οι μετρητές διαφέρουν ανάλογα με την έκδοση του πρωτοκόλλου. Στην πρώτη έκδοση, που έτυχε ευρείας υλοποίησης από κατασκευαστές και ονομάστηκε OpenFlow Switch Specification 1.0.0 [OFspec10], τα πεδία ταύτισης είναι συνολικά 12 (**Πίνακας 2-1).** Όπως παρατηρούμε τα πεδία που μπορούν να χρησιμοποιηθούν ως κριτήριο για την ταύτιση των πακέτων, αν εξαιρέσουμε την πόρτα εισόδου, αφορούν το Επίπεδο 2, Επίπεδο 3 και 4 της στοίβας πρωτοκόλλων.

Πεδίο	Bits	Εφαρμογή	Περιγραφή
Ingress Port	Implementation depedent	All packets	Numerical representation of incoming port starting at 1
Ethernet source address	48	All packets on enabled ports	
Ethernet destination address	48	All packets on enabled ports	
Ethernet type	16	All packets on enabled ports	OF switch is required to match the type in standard Ethernet and 802.2 with a SNAP header and OUI of 0x000000. The special value of 0x05FF is used to match all 802.3 packets without SNAP headers.
VLAN id	12	All packets of Ethernet type 0x8100	
VLAN priority	3	All packets of Ethernet type 0x8100	VLAN PCP field
IP source address	32	All IP and ARP packets	Can be subnet masked
IP destination address	32	All IP and ARP packets	Can be subnet masked
IP protocol	8	All IP and IP over Ethernet , ARP packets	Only the lower 8 bits of the ARP op-code are used
IP ToS bits	6	All IP packets	Specify as 8-bit value and place ToS in upper 6 bits.
Transport source port / ICMP Type	16	All TCP, UDP, and ICMP packets	Only lower 8 bits used for ICMP Type
Transport destination port / ICMP Code	16	All TCP, UDP, and ICMP packets	Only lower 8 bits used for ICMP Code

Πίνακας 2-1: Πεδία ταύτισης του OpenFlow Specification 1.0.0

Στην έκδοση OpenFlow Switch Specification 1.4.0 [OFspec14] έχει ορισθεί ένας πολύ μεγαλύτερος αριθμός πεδίων ταύτισης (**Πίνακας 2-2)**. Όπως παρατηρούμε τα πεδία που μπορούν να χρησιμοποιηθούν ως κριτήριο για την ταύτιση των πακέτων, αν εξαιρέσουμε την πόρτα εισόδου, αφορούν το Επίπεδο 2, Επιπέδο 2.5 (Multi-Protocol Label Switching - MPLS), Επίπεδο 3 και 4 της στοίβας πρωτοκόλλων. Με έντονο χρώμα σημειώνονται τα υποχρεωτικά πεδία ταύτισης που πρέπει να υποστηρίζει η κάθε συμβατή συσκευή με τη συγκεκριμένη έκδοση του πρωτοκόλλου. Στην στήλη με τίτλο «Περιγραφή» αναφέρονται οι προϋποθέσεις που πρέπει να πληρεί ένα πλαίσιο/πακέτο για να μπορεί να γίνει η αντιστοίχιση πεδίων επικεφαλίδων του με τα πεδία ταύτισης, όπως αυτά ορίζονται μέσα στους πίνακες προώθησης.

Πεδίο	**Bits**	**Περιγραφή**
Switch input port	***32***	***None***
Switch physical input port	32	IN_PORT present
Metadata passed between tables	64	None
Ethernet destination address	***48***	***None***
Ethernet source address	***48***	***None***
Ethernet frame type	***16***	***None***
VLAN id	12+1	None
VLAN priority	3	VLAN_VID!=NONE
IP DSCP (6 bits in ToS field)	6	ETH_TYPE=0x800 or ETH_TYPE=0x86dd
IP ECN (2 bits in ToS field)	2	ETH_TYPE=0x0800 or ETH_TYPE=0x86dd
IP protocol	***8***	***ETH_TYPE=0x0800 or ETH_TYPE=0x86dd***
IPv4 source address	***32***	***ETH_TYPE=0x0800***
IPv4 destination address	***32***	***ETH_TYPE=0x0800***
TCP source port	***16***	***IP_PROTO=6***
TCP destination port	***16***	***IP_PROTO=6***
UDP source port	***16***	***IP_PROTO=17***
UDP destination port	***16***	***IP_PROTO=17***
SCTP source port	8	IP_PROTO=132
SCTP destination port	8	IP_PROTO=132
ICMP type	8	IP_PROTO=1
ICMP code	8	IP_PROTO=1
ARP opcode	16	ETH_TYPE=0x0806
ARP source IPv4 address	32	ETH_TYPE=0x0806
ARP target IPv4 address	32	ETH_TYPE=0x0806
ARP source hardware address	48	ETH_TYPE=0x0806
ARP target hardware address	48	ETH_TYPE=0x0806
IPv6 source address	***128***	***ETH_TYPE=0x86dd***
IPv6 destination address	***128***	***ETH_TYPE=0x86dd***
IPv6 Flow Label	20	ETH_TYPE=0x86dd
ICMPv6 type	8	IP_PROTO=58
ICMPv6 code	8	IP_PROTO=58
Target address for ND	128	ICMPV6_TYPE=135 or ICMP6_TYPE=136
Source link-layer for ND	48	ICMPV6_TYPE=135
Target link-layer for ND	48	ICMP6_TYPE=136

MPLS label	20	ETH_TYPE=0x8847 or ETH_TYPE=0x8848
MPLS TC	3	ETH_TYPE=0x8847 or ETH_TYPE=0x8848
MPLS BoS bit	1	ETH_TYPE=0x8847 or ETH_TYPE=0x8848
PBB I-SID	24	ETH_TYPE=0x88E7
Logical Port Metadata	64	None
IPv6 Extension Header pseudo-field	9	ETH_TYPE=0x86dd
PBB UCA header field	1	ETH_TYPE=0x88E7

Πίνακας 2-2: Πεδία ταύτισης του OpenFlow Specification 1.4.0

Ένας OpenFlow Controller μέσω του OpenFlow protocol μπορεί να εισάγει, αφαιρεί, ανανεώνει flow entries που βρίσκονται σε ένα διασυνδεδεμένο με αυτόν OpenFlow switch. Υπάρχει η δυνατότητα εισαγωγής νέων κανόνων προληπτικά (proactively), πριν την άφιξη δηλαδή κάποιου πακέτου που ταιριάζει στο εκάστοτε flow entry. Δίνεται επίσης η δυνατότητα χειρισμού ενός πακέτου που δεν αντιστοιχεί με κάποιον ήδη εγκαθιδρυμένο κανόνα, αφού το OpenFlow Switch έχει τη δυνατότητα να αποθηκεύσει προσωρινά ένα πακέτο και να ρωτήσει τον OpenFlow Controller για τον τρόπο χειρισμού του πακέτου. Η ερώτηση αυτή ουσιαστικά απαντάται με την εγκαθίδρυση ενός καινούριου OpenFlow entry από πλευράς OpenFlow Controller στο switch.

Τα Actions, που περιλαμβάνονται ως μέρος του κάθε Flow entry, δύνανται να υπαγορεύουν την έξοδο του πακέτου από μια πόρτα (forwarding), την απόρριψή του (dropping), την τροποποίηση (modification) κάποιων επικεφαλίδων πρωτοκόλλων που λαμβάνει υπόψη του το OpenFlow ως πεδία ταύτισης ή ακόμα και την μεταφορά της απόφασης σε κάποιο άλλο Flow table που ακολουθεί στην αλυσίδα των Flow tables.

3 Εικονικοποίηση

Η έννοια της εικονικοποίησης έχει χρησιμοποιηθεί στον σχεδιασμό υπολογιστικών συστημάτων και κυρίως συστημάτων μνήμης από την δεκαετία του 1960 [McGe65]. Ο σχεδιασμός των σύγχρονων υπολογιστικών συστημάτων χρησιμοποιεί εκτεταμένα τεχνολογίες εικονικοποίησης, με σκοπό την αποσύζευξη (decoupling) των μοντέλων παροχής υπηρεσιών από τους φυσικούς πόρους που χρησιμοποιούνται για την υλοποίηση των υπηρεσιών [Casa10].

Δυο χαρακτηριστικά παραδείγματα εικονικοποίησης υπολογιστικών και δικτυακών πόρων είναι αυτό των εικονικών μηχανών (virtual machines) και των εικονικών ιδιωτικών δικτύων (Virtual Private Networks- VPN) αντίστοιχα. Η εντατική ανάπτυξη αντίστοιχων τεχνολογιών εικονικοποίησης για ένα μεγάλο εύρος υπηρεσιών και υπολογιστικών πόρων, όπως είναι οι δικτυακοί πόροι και τα μέσα αποθήκευσης, είχε ως αποτέλεσμα την δημιουργία νέων στρωμάτων αφαίρεσης (abstraction layers).

Οι υπολογιστικοί πόροι και τα δεδομένα με την βοήθεια των νέων στρωμάτων αφαίρεσης είναι δυνατόν να μην είναι συσχετισμένα με συγκεκριμένους φυσικούς πόρους, δίνοντας με αυτό τον τρόπο νέες διαχειριστικές ευκολίες όπως είναι η ελαστικότητα (elasticity), η μετανάστευση (migration) και η δημιουργία νέων καινοτόμων υπηρεσιών (π.χ νεφο-υπολογιστική – cloud computing) που αναλύονται στη συνέχεια.

Στο τομέα των υπολογιστικών συστημάτων (π.χ. δομές μνήμης και συστήματα αποθήκευσης), οι τεχνικές εικονικοποίησης υιοθετήθηκαν εκτεταμένα πολύ πριν αναπτυχθούν αντίστοιχες τεχνικές στο κόσμο της δικτύωσης, με αποτέλεσμα σε πολλές περιπτώσεις η δικτύωση των υπολογιστικών συστημάτων να γίνεται τροχοπέδη στην συνολική λειτουργία μεγάλων συστημάτων όπως είναι τα κέντρα δεδομένων (datacenter) [Hami09].

3.1 Εικονικοποίηση Δικτύων Υπολογιστών – Network Virtualization

Η βασική λειτουργία της δρομολόγησης πακέτων μπορεί να υλοποιηθεί σε τυχαίες τοπολογίες. Ωστόσο, ένα μεγάλο μέρος των δικτυακών υπηρεσιών, όπως είναι η παροχή συγκεκριμένης ποιότητας υπηρεσίας (Quality of Service – QoS) στην επικοινωνία, η παροχή λιστών πρόσβασης με σύνθετα κριτήρια (Access Control Lists –ACL) και η πολιτικοπαγής δρομολόγηση (policy-based routing), εξαρτάται από την τοπολογία (topology) [Casa07], [Ioan00], [Wang08]. Η λειτουργία τέτοιων υπηρεσιών είναι δύσκολη, χρονοβόρος και ευάλωτη σε λάθη.

Για τον παραπάνω λόγο, οι τεχνολογίες εικονικοποίησης δικτύων (network virtualization) αναπτύχθηκαν με σκοπό την καλύτερη εκμετάλλευση των δικτυακών υποδομών καθώς και την πιο ευέλικτη και αποτελεσματική διαχείρισή τους. Οι

συγκεκριμένες τεχνολογίες δίνουν πλήθος νέων δυνατοτήτων μέσω της ταυτόχρονης χρήσης των φυσικών και λογικών δικτυακών πόρων από πολλαπλούς χρήστες ή/και την συνάθροιση πόρων που οδηγεί σε αυξημένες επιδόσεις. Σε συνδυασμό με τις τεχνολογίες εικονικοποίησης εξυπηρετητών, συμπληρώνουν το μωσαϊκό των τεχνολογιών εικονικοποίησης υπολογιστικών συστημάτων.

Η ιδέα της εικονικοποίησης συναντάται σε πολλά ήδη καθιερωμένα και ευρέως χρησιμοποιούμενα πρωτόκολλα της στοίβας δικτυακών πρωτοκόλλων. Έχει χρησιμοποιηθεί στο παρελθόν για να αυξήσει τον **δείκτη χρησιμοποίησης των πόρων** (utilization), για τον **λογικό διαχωρισμό της κίνησης** (logical separation) μεταξύ διαφορετικών οντοτήτων, για την **απλοποίηση της διαχείρισης** (network management) καθώς και για την **παροχή ασφαλούς διασύνδεσης** (security) πάνω από αναξιόπιστα δίκτυα. Ακολουθούν χαρακτηριστικά παραδείγματα χρήσης και η αναφορά σε συγκεκριμένα πρωτόκολλα.

3.1.1 Πρώιμες Υλοποιήσεις Εικονικοποίησης Δικτύων Υπολογιστών

Οι υλοποιήσεις εικονικοποίησης, που τυγχάνουν ευρείας αποδοχής, περιλαμβάνουν και πρωτόκολλα επικοινωνίας (IEEE 802.1Q, MPLS) που υλοποιούνται σε επίπεδο ζεύξεως/διεπαφών (link/interfaces) καθώς και σε επίπεδο δικτυακών κόμβων, είτε αυτοί είναι μεταγωγείς (switches) είτε δρομολογητές (routers).

Η παροχή από άκρο-σε-άκρο (end-to-end) ή πολυσημειακής-σε-πολυσημειακή (multipoint-to-multipoint) διασύνδεσης είναι δυνατή, με χρήση πολλαπλών τεχνικών εικονικής ιδιωτικής δικτύωσης (Virtual Private Networking- VPN). Οι τεχνικές εικονικής ιδεατής δικτύωσης όσον αφορά ζεύξεις στο επίπεδο 2 της στοίβας περιλαμβάνουν λύσεις που στηρίζονται σε πρωτόκολλα που ανήκουν στο επίπεδο 2 μέχρι και το επίπεδο 5 της στοίβας πρωτοκόλλων. Οι διάφορες τεχνικές εξασφαλίζουν διασύνδεση των άκρων μέσω της ενθυλάκωσης μονάδων δεδομένων πρωτοκόλλων (Payload Data Unit–PDU) σε μονάδες δεδομένων υπηρεσίας (Service Data Unit – SDU) επιπέδου 2 έως 5.Στην περίπτωση που παρέχεται από το μηχανισμό ενθυλάκωσης και η δυνατότητα κρυπτογράφησης της πληροφορίας (encryption)

παρέχεται εκτός από την υπηρεσία διασύνδεσης και μυστικότητα στην μεταφορά των δεδομένων, όπως για παράδειγμα συμβαίνει κατά τη χρήση του πρωτοκόλλου IPSEC.

Για παράδειγμα, το Generic Routing Protocol (GRE) λειτουργεί ενθυλακώνοντας πακέτα IP (IP packets) ή πλαίσια Ethernet (Ethernet frames) σε IP πακέτα. Από την πλευρά του χρήστη γίνεται αντιληπτή μόνο η ύπαρξη των IP πακέτων ή των Ethernet frames στα ακραία σημεία εξόδου/εισόδου τους, χωρίς να γνωρίζει τις λεπτομέρειες ενθυλάκωσης οι οποίες κάνουν δυνατή την αποστολή της πληροφορίας μεταξύ των 2 άκρων που ενώνει ένα GRE tunnel. Ο χρήστης μάλιστα της εικονικής υπηρεσίας δικτύωσης δεν έχει κανένα έλεγχο για τον τρόπο που έχει υλοποιηθεί η υπηρεσία διασύνδεσης των άκρων. Ο διαχωρισμός της κίνησης των χρηστών και η δυνατότητα ταυτόχρονης εξυπηρέτησής τους αυξάνει την χρησιμοποίηση των δικτυακών πόρων (ζεύξεων, δρομολογητών, μεταγωγέων) στην πλευρά του παρόχου.

Συμπληρωματικά των εικονικών ζεύξεων και με την χρήση πρωτοκόλλων όπως το IEEE 802.1Q ή τεχνολογιών Virtual Routing and Forwarding (VRF) σε επίπεδο δικτύου (IP layer) μπορεί να γίνει διαφορετικός χειρισμός της προωθούμενης κίνησης πάνω στις δικτυακές συσκευές ανά πελάτη . Ο διαχωρισμός, σε αυτές τις περιπτώσεις, γίνεται συνήθως με κριτήριο κάποια ετικέτα, όπως για παράδειγμα το VLAN tag στην περίπτωση του IEEE 802.1Q ή με κριτήριο τη διεπαφή εισόδου της κίνησης (ingress interface).

3.1.2 Σχεδιαστικές Ελλείψεις Πρώιμων Υλοποιήσεων Εικονικοποίησης

Το στρώμα αφαίρεσης δικτύου αποτελεί αναγκαία προϋπόθεση για την ανάπτυξη ενός λογισμικού ελέγχου που θα είναι ανεξάρτητο της τοπολογίας και θα δίνει την δυνατότητα παραμετροποίησης και ελέγχου των δικτυακών συσκευών ως ένα ενιαίο σύνολο.

Η ύπαρξη ανόμοιων τεχνολογιών και πρωτοκόλλων που δομήθηκαν με την προοπτική της κοινής χρήσης δικτυακών πόρων (shared resources), λύνοντας επιμέρους προβλήματα, δεν είναι αρκετή για να δομηθεί ένα ενιαίο γενικευμένο στρώμα αφαίρεσης δικτύου.

Επιμέρους υλοποιήσεις στρωμάτων αφαίρεσης δικτύων έχουν σχεδιασθεί για την κάλυψη αναγκών που αφορούν συγκεκριμένες δικτυακές υπηρεσίες, όπως είναι η εφαρμογή πολιτικών ασφαλείας σε μεγάλα δίκτυα [Ioan00] και η παροχή ποιότητας υπηρεσίας σε ασύρματα τοπικά δίκτυα (Wireless Local Area Network –WLAN) [Cisco].

3.1.3 Η Εξέλιξη των Εικονικών Δικτύων

Τα εικονικά δίκτυα προάγουν την ιδέα της συνεκμετάλλευσης, που εισήγαγε η στατιστική πολυπλεξία πακέτων στα δίκτυα μεταγωγής πακέτων (packet-switched networks) και κυρίως των IP δικτύων. Ο συνδυασμός δικτυακών πόρων και η λογική συνάθροισής τους, υπό τον έλεγχο διαφορετικών διαχειριστικών οντοτήτων, δημιουργεί τις προϋποθέσεις για την κατασκευή εικονικών δικτύων ελεγχόμενων από διαφορετικούς χρήστες.

Τα εικονικά δίκτυα δομούνται με τέτοιο τρόπο ώστε να μπορούν να προσφέρουν τυπικές υπηρεσίες που συναντώνται σε κλασσικά δικτυακά περιβάλλοντα, όπως είναι η διασύνδεση σε Επίπεδο 2 και 3 της στοίβας (Layer 2 και 3) και η υπηρεσία ονομάτων τομέων (Domain Name Service - DNS) καθώς και η υπηρεσία πρωτοκόλλου δυναμικής διάρθρωσης κόμβων (Dynamic Host Configuration Protocol - DHCP). Θα μπορούσαμε να συνοψίσουμε τις βασικές λειτουργίες των υποδομών, που προσφέρουν υπηρεσίες εικονικοποίησης δικτύων στα ακόλουθα [Argy13]:

- **Πολυπλεξία (multiplexing/concurrence)**

  Ταυτόχρονη χρήση κοινών δικτυακών πόρων υποδομής (infrastructure resources) από πολλαπλά εικονικά δίκτυα και κατ΄ επέκταση πολλαπλούς χρήστες.

- **Απομόνωση (Isolation)**

  Διαχωρισμός των λειτουργιών και κατανομή των πόρων οι οποίοι είναι δυνατόν να χρησιμοποιηθούν από κάθε εικονικό δίκτυο, ώστε να μπορούν να δίνονται εγγυήσεις όσον αφορά την ιδιωτικότητα (privacy) και το επίπεδο της παρεχόμενης υπηρεσίας (Quality of Service – QoS) ανά εικονικό δίκτυο/χρήστη.

- **Αφαίρεση (Abstraction)**

  Η ταυτόχρονη χρήση δικτυακών πόρων από πολλαπλά εικονικά δίκτυα καθώς και η συνάθροισή τους απαιτεί την δημιουργία ενός στρώματος αφαίρεσης δικτύου (network abstraction layer). Από την πλευρά του χρήστη γίνεται αντιληπτή η ύπαρξη μόνο του υποσυνόλου των πόρων που του έχουν ανατεθεί, ενώ από πλευράς του παρόχου το συγκεκριμένο στρώμα αφαίρεσης δικτύου επιτρέπει τον σαφή διαχωρισμό των πόρων του κάθε χρήστη.

- **Ελαστικότητα (Elasticity)**

  Η ευέλικτη πρόσθεση/αφαίρεση δικτυακών πόρων που αποδίδονται σε εικονικά δίκτυα επιτρέποντας την εύκολη προσαρμογή στις δυναμικές ανάγκες των χρηστών.

- **Μετανάστευση (Migration)**

  Η ανακατανομή των πόρων που χρησιμοποιούνται και η αναδόμηση αυθαίρετων τοπολογιών πάνω από τη φυσική δικτυακή υποδομή για διαχειριστικούς λόγους, όπως για παράδειγμα η ισοκατανομή φορτίου (load-balancing) ή αντιμετώπιση αστοχιών (fail-over).

- **Δυνατότητα Προγραμματισμού (Programmability)**

  Ικανότητα εφαρμογής και χρήσης προγραμματιστικών διαδικασιών και εργαλείων για την διαχείριση διαφορετικών πρωτοκόλλων και δομών δικτύωσης. Η δυνατότητα δυναμικού προγραμματισμού της λειτουργίας των δικτυακών συσκευών μπορεί να προσφέρει εύκολο σχεδιασμό νέων υπηρεσιών για τους χρήστες.

4 Διασύνδεση Εικονικών Υπολογιστικών Πόρων στο Επίπεδο Προώθησης

Network Virtualization over Heterogeneous Federated Infrastructures: Data Plane Connectivity

4.1 Ερευνητικός Στόχος

Οι αρχιτεκτονικές ομόσπονδων πειραματικών υποδομών, όπως είναι το Slice-based Federation Architecture (SFA) παρέχουν την δυνατότητα επικοινωνίας σε επίπεδο διαχείρισης των δικτυακών και υπολογιστικών πόρων, αλλά δεν δίνουν λύση στο πρόβλημα διασύνδεσης σε επίπεδο προώθησης δεδομένων.

Η χρήση μάλιστα εικονικών πόρων (virtual resources) σε ένα ομόσπονδο δικτυακό περιβάλλον απαιτεί τεχνικές δημιουργίας εικονικών δικτύων σε όλο το εύρος της ομοσπονδίας. Τα εικονικά δίκτυα πρέπει να παρέχουν τυπικές δικτυακές υπηρεσίες που μπορούμε να βρούμε σε ένα συνηθισμένο περιβάλλον, όπως είναι η διασύνδεση σε επίπεδο μέσου (data link layer) και σε δικτυακό επίπεδο (network layer). Επίσης θα πρέπει να έχουν πρόσθετα χαρακτηριστικά, όπως είναι η ελαστικότητα στον τρόπο παροχής των πόρων (elasticity), δυνατότητα αναδιάταξης των πόρων και δημιουργία πειραματικών τοπολογιών ανεξαρτήτως της φυσικής τοπολογίας της φυσικής υποδομής.

Στην περίπτωση μάλιστα που οι τεχνολογίες που χρησιμοποιούνται για την υλοποίηση του επιπέδου προώθησης των ομόσπονδων υποδομών είναι διαφορετικές, αυξάνει η δυσκολία, δεδομένης της απαιτούμενης ταυτόχρονης διαχείρισης και παραμετροποίησης ανομοιογενών πόρων που ανήκουν στις ομόσπονδες διαχειριστικές οντότητες (federated domains).

Στόχος μας είναι η δικτυακή διασύνδεση μιας εικονικής υπολογιστικής οντότητας (virtual node) με μια άλλη αντίστοιχη οντότητα η οποία βρίσκεται σε κάποια ομόσπονδη διαχειριστική περιοχή (federated domain).

Η ερευνητική μας εργασία εστιάζει στην μελέτη και δημιουργία ενός μηχανισμού διασύνδεσης ετερογενών ομόσπονδων υποδομών εικονικοποιημένων πόρων για την ταυτόχρονη εξυπηρέτηση πολλαπλών χρηστών. Η λύση που καλούμαστε να δώσουμε θα πρέπει να προσφέρει τα χαρακτηριστικά εικονικής δικτύωσης, όπως αυτά έχουν περιγραφεί στην Παράγραφο 3.1.3

4.2 Αρχές Σχεδίασης

Το κύριο αντικείμενο της εργασίας μας, αναφορικά με το επίπεδο προώθησης δεδομένων ετερογενών υποδομών, είναι η συνένωση των εικονικών δικτύων που δημιουργούν οι ερευνητές σε ετερογενείς ομόσπονδες υποδομές δικτυακής εικονικοποίησης.

Καθιστώντας λειτουργικά τα εικονικά δίκτυα διαφορετικών υποδομών ως μια ενιαία οντότητα, ικανοποιούμε την απαίτηση ομοσπονδοποίησης των υποδομών σε επίπεδο προώθησης δεδομένων. Η προσέγγιση μας προσφέρει τη σχεδίαση και την υλοποίηση τεχνικών που μπορούν να προσφέρουν την ευρεία χρήση της λύσης που προτείνουμε σε πληθώρα εφαρμογών δικτυακής εικονικοποίησης.

Τα εικονικά δίκτυα αποτελούνται από εικονικές διεπαφές (virtual interfaces), εικονικές ζεύξεις (virtual links), εικονικούς δρομολογητές (virtual routers) και εικονικούς μεταγωγείς (virtual switches). Η περιγραφή της συνένωσης καθώς και οι τεχνολογίες που χρησιμοποιήσαμε αφορούν δίκτυα ικανά να μεταφέρουν Ethernet frames. Ως βάση για την ανάπτυξη του πρωτοτύπου διασύνδεσης χρησιμοποιήσαμε ένα μεταγωγέα Ethernet υλοποιημένο σε λογισμικό που ονομάζεται Open vSwitch [OVS]. Το συγκεκριμένο λογισμικό έχει αναπτυχθεί από την ερευνητική κοινότητα ως ανοιχτό λογισμικό και μάλιστα μέσα στο 2012 υιοθετήθηκε ως κομμάτι του λειτουργικού συστήματος Linux.

Οι μέθοδοι που αναπτύξαμε καθώς και η προτυποποίηση έγινε με χρήση πραγματικών υποδομών που προσφέρουν υπηρεσίες σε ερευνητές. Το περιβάλλον που επιλέξαμε θα έπρεπε να πληροί και την προϋπόθεση της ετερογένειας μεταξύ των υποδομών που θα έπρεπε να συνενωθούν, καθώς και να υποστηρίζουν ως βασική δικτυακή λειτουργία την προώθηση Ethernet frames. Για τους παραπάνω λόγους το καλύτερο περιβάλλον κρίθηκε αυτό του ερευνητικού προγράμματος Network Innovation over Virtualized Infrastructures (NOVI).

Η μεθοδολογία που ακολουθήσαμε προσφέρει τα χαρακτηριστικά εικονικής δικτύωσης, όπως αυτά έχουν περιγραφεί στην Παράγραφο 3.1.3. Τα κύρια σημεία της σχεδίασης που ακολουθήσαμε είναι τα ακόλουθα:

- Δυνατότητα διασύνδεσης ετερογενών υποδομών εικονικοποίησης στο επίπεδο προώθησης δεδομένων.

  Συμβατός σχεδιασμός με υπάρχουσες ετερογενείς υποδομές και να μπορεί να υποστηρίξει την ανάπτυξη πρωτοτύπου για πλατφόρμες όπως είναι το PlanetLab, το VINI και το FEDERICA, που αποτελούν πειραματικές πλατφόρμες μεγάλης κλίμακας.

- Δυνατότητα εικονικοποίησης που ικανοποιεί την απαίτηση για διασύνδεση σε επίπεδο Ethernet frames (Layer 2) ή/και σε επίπεδο σε IP packets (Layer 3).

  Είναι δυνατή η διασύνδεση υποδομών που υλοποιούν την εικονικοποίηση των δικτυακών ζεύξεων σε Layer 3 με υποδομές που την υλοποιούν σε Layer 2 (π.χ. PlanetLab –FEDERICA).

- Δυνατότητα παραμετροποίησης συνδέσεων πολλαπλών σημείων που βρίσκονται σε πολλαπλές διαχειριστικές περιοχές.

  Η δυνατότητα διασύνδεσης του επιπέδου προώθησης δεδομένων των ομόσπονδων υποδομών σε πολλαπλά σημεία απαιτείται για να είναι εφικτή η δημιουργία πολύπλοκων τοπολογιών και σύνθετων πειραμάτων πάνω σε αυτές (π.χ. load-balancing, fail-over).

- Χρήση του πρωτοκόλλου Generic Routing Encapsulation(GRE) και του IEEE 802.1Q.

Η ευρεία χρήση του GRE πρωτοκόλλου καθώς και η δυνατότητα ενθυλάκωσης Ethernet frames καθώς και IP πακέτων πάνω από IP δίκτυα, σε συνδυασμό με την ευρεία χρήση του IEEE 802.1Q κάνουν την λύση μας συμβατή με πληθώρα υποδομών. Τα συγκεκριμένα μάλιστα πρωτόκολλα έχουν εγγενή δυνατότητα διαχωρισμού ροών, ικανότητα αναγκαία για την διαδικασία της εικονικοποίησης των ζεύξεων όπως θα εξηγηθεί αναλυτικότερα παρακάτω.

- Χρήση προγραμματιζόμενου σε επίπεδο διαχείρισης και σε επίπεδο ελέγχου μεταγωγέα υλοποιημένο σε λογισμικό.

  Το Open vSwitch που χρησιμοποιήσαμε έχει την ικανότητα να διασυνδέει φυσικές διεπαφές (π.χ Ethernet interfaces) και λογικές διεπαφές (TUN/TAP, GRE over IP) μέσω του API που διαθέτει στο επίπεδος διαχείρισης. Επίσης είναι ικανό να προγραμματίζεται στο επίπεδο ελέγχου από το OpenFlow, πράγμα που δίνει δυνατότητα κεντρικοποιημένου ελέγχου προώθησης ροών, όπως έχει περιγραφεί στην Παράγραφο 2.4.

4.3 Το Ερευνητικό Οικοσύστημα του NOVI

Το ερευνητικό πρόγραμμα NOVI ανέπτυξε τους απαιτούμενους αλγορίθμους, τα διαχειριστικά εργαλεία, και το μοντέλο πληροφοριών (information model) που απαιτείται για την συνένωση ετερογενών ομόσπονδων υποδομών, ώστε να είναι δυνατή η εύρεση, αποτύπωση, παρακολούθηση και παροχή των εικονικών δικτυακών πόρων σε συνδυασμό με τους κατανεμημένους υπολογιστικούς πόρους και μέσα αποθήκευσης. Το περιβάλλον που χρησιμοποίησε για τις δοκιμές του συνολικού πλαισίου και το οποίο έγινε αντικείμενο πειραματισμού και για τα δικά μας πρωτότυπα είναι ένα ιδιωτικό αντίγραφο της υποδομής του PlanetLab και η υποδομή του FEDERICA.

Καμία υποδομή πριν την ανάπτυξη του NOVI δεν είχε την ικανότητα ταυτόχρονης διαχείρισης και παραμετροποίησης ανομοιογενών πόρων που ανήκουν σε ομόσπονδες διαχειριστικές οντότητες (federated domains) και ούτε υπήρχε κάποια μεθοδολογία διασύνδεσης των επιπέδων προώθησης δεδομένων των ομόσπονδων διαχειριστικών οντοτήτων που μελετούμε στην εργασία μας.

Το ιδιωτικό αντίγραφο του PlanetLab, που χρησιμοποιήθηκε στο NOVI είναι μικρότερης κλίμακας και χρησιμοποιεί το Internet ως δικτυακό υπόστρωμα για την διασύνδεση των υπολογιστικών πόρων, όπως ακριβώς και το PlanetLab. Οι υπολογιστικοί πόροι του ιδιωτικού PlanetLab που χρησιμοποιήθηκε για το σύνολο των δοκιμών μας ήταν κατανεμημένοι σε διάφορα σημεία παρουσίας (Point of Presence – PoPs) ανά την Ευρώπη και ελέγχονταν από μια κεντρικοποιημένη μονάδα διαχείρισης (myPLC) σε ένα από τα σημεία παρουσίας.

Το FEDERICA χρησιμοποιεί, ως υπόστρωμα δικτυακής διασύνδεσης των κόμβων του, δικτυακούς πόρους από το GÉANT [GEANT]. Οι δικτυακοί πόροι είναι αποκλειστικής δεσμεύσεως (bandwidth reservation) για τα πειράματα που εκτελούνται πάνω στο FEDERICA. Οι υπολογιστικοί πόροι του FEDERICA είναι επίσης κατανεμημένοι ανά την Ευρώπη. Το επίπεδο διαχείρισης του NOVI είχε πρόσβαση με αυξημένα δικαιώματα, μέσω API, στο επίπεδο διαχείρισης του FEDERICA.

Μια διαφορετική προσέγγιση εικονικοποίησης από το PlanetLab ακολουθήθηκε στο FEDERICA. Το FEDERICA ήθελε να εξασφαλίσει ελεγχόμενες συνθήκες στο περιβάλλον που έτρεχαν τα πειράματα. Αυτός ήταν ο λόγος που αντί να χρησιμοποιήσει το Internet για την διασύνδεση των κατανεμημένων σημείων παρουσίας των πόρων ανά την Ευρώπη, έκανε χρήση λογικών δρομολογητών (logical routers) που υλοποιούνταν πάνω σε εξοπλισμό της Juniper και κυκλωμάτων SDH/SONET ταχύτητας 1Gbps. Οι φυσικοί δρομολογητές ήταν Juniper MX-480 [JUNOS] και οι μεταγωγείς EX-4200, σε συνδυασμό με τα SDH/SONET κυκλώματα, αποτελούσαν τους δικτυακούς πόρους που διαμοιράζονταν. Για την εικονικοποίηση των υπολογιστικών πόρων χρησιμοποιήθηκε η πλατφόρμα ESXi της VMware. Το επίπεδο διαχείρισης του FEDERICA είναι ικανό να διαμερίζει τους διαφόρους φυσικούς πόρους που αποτελούν την υποδομή, εκμεταλλευόμενο τις ικανότητες εικονικοποίησης που προσφέρουν οι επιμέρους συσκευές μέσω των API τους (π.χ logical routers πάνω σε MX-480).

4.4 Ανασκόπηση της Υπάρχουσας Κατάστασης

Ένα από τα πρώτα συνεργατικά, μεγάλης κλίμακας, περιβάλλοντα πειραματισμού που αναπτύχθηκε ήταν το PlanetLab [PlanetLabH]. Ο αρχικός σκοπός ανάπτυξής του ήταν η δοκιμή πολλαπλών υπολογιστικών και δικτυακών υπηρεσιών διαφορετικών ερευνητών/χρηστών, σε ένα περιβάλλον που θα επέτρεπε τον διαχωρισμό τους, δημιουργώντας συνθήκες απομόνωσης (isolation) [Chun03].

Το σύνολο των διαμοιραζόμενων πόρων, που είχε πρόσβαση και δυνατότητα χρήσης ο κάθε ερευνητής, ήταν κατανεμημένο σε ένα υποσύνολο της συνολικής υποδομής και ονομαζόταν **slice**. Οι κατανεμημένοι υπολογιστικοί πόροι που αποτελούσαν το slice ονομάζονται **slivers** και είχαν την ικανότητα να επικοινωνούν μεταξύ τους με χρήση του Internet, χωρίς να δίνεται ως εκ τούτου καμία εγγύηση για την ποιότητα της δικτυακής υπηρεσίας. Το PlanetLab μέσω του δομικού στοιχείου (module) VNET [Huan05] μπορούσε να εξασφαλίσει τον διαχωρισμό μεταξύ των δικτυακών πόρων που χρησιμοποιούσαν οι πολλαπλοί χρήστες των κόμβων, δίνοντας έτσι στο sliver κάθε χρήστη την δυνατότητα (περιορισμένης) χρήσης raw IP sockets. Το δομικό στοιχείο VNET περιόριζε τους πειραματισμούς του χρήστη πάνω από το δικτυακό επίπεδο του πρωτοκόλλου IP (IP network layer), χωρίς να δίνει την δυνατότητα πειραματισμών με Ethernet frames ή με υπερκείμενες τοπολογίες στο επίπεδο δικτύου δικής του επιλογής.

Η ικανότητα δημιουργίας ελεγχόμενων από τον χρήστη δικτυακών πόρων έγινε αργότερα με το πλαίσιο που αναπτύχθηκε στην ερευνητική υποδομή VINI. Οι δικτυακές δυνατότητες του κάθε κόμβου του PlanetLab (PL node) επεκτάθηκαν στο VINI (PL-VINI node), δίνοντας νέες δυνατότητες στα sliver τα οποία δημιουργούνταν στους PL-VINI κόμβους. Μια τροποποιημένη έκδοση του δομικού στοιχείου (i) TUN/TAP του πυρήνα Linux, (ii) ο δρομολογητής λογισμικού (software router) Click, (iii) η σουίτα πρωτοκόλλων δρομολόγησης XORP [Hand05], (iv) το πακέτο λογισμικού OpenVPN [OpenVPN] και η δυνατότητα δημιουργίας UDP tunnels για την δημιουργία point-to-point ζεύξεων υπερκείμενων δικτύων εμπλούτισαν τις δυνατότητες δικτυακής συνδεσιμότητας των slices. Τα παραπάνω στοιχεία που προσέθεσε το VINI στην υποδομή του, σε σχέση με τον πρόγονό του (PlanetLab),

έκαναν δυνατό τον πειραματισμό πάνω σε πρωτόκολλα δρομολόγησης, προώθησης πακέτων δημιουργώντας μια υποδομή δικτυακής εικονικοποίησης [Bavie06].

Περαιτέρω βελτιώσεις σε επίπεδο πυρήνα λειτουργικού συστήματος και εργαλείων υπό την ονομασία Trellis υλοποιήθηκαν για το περιβάλλον του VINI [Bhat08]. Το Trellis βασίστηκε σε δυο container-based τεχνολογίες εικονικοποίησης: (i) τον VServer [Vserver], το NetNS [LXC] και σε ένα (iii) μηχανισμό Ethernet over GRE [RFC2784], με σκοπό να παρέχει υψηλής ταχύτητας δυνατότητες εικονικοποίησης στη πλατφόρμα του VINI.

Το ProtoGeni αποτελεί μια προσπάθεια ενοποίησης μεγάλης κλίμακας υποδομών που ελέγχονται από το GENI που ακολούθησε. Η δικτυακή διασύνδεση των υποδομών του ProtoGeni υποστηρίζεται από το δίκτυο κορμού του Internet2 και έχουν γίνει οι κατάλληλες παραμετροποιήσεις ώστε να μπορεί να τεμαχιστεί το σύνολο των δικτυακών πόρων που έχουν αποδοθεί από πλευράς του Internet2, ώστε να ικανοποιούνται οι ανάγκες πολλαπλών πειραμάτων που τρέχουν πάνω από τις ενοποιημένες υποδομές του ProtoGeni. Έχουν χρησιμοποιηθεί τεχνολογίες και υποδομές από το PlanetLab, το VINI καθώς και το Emulab [Hible08]. Η υποδομή του ProtoGeni επιτρέπει την δημιουργία τοπολογιών σε επίπεδο προώθησης Ethernet frames (Layer 2), διαχωρίζοντας την κίνηση των πολλαπλών πειραμάτων με χρήση του πρωτοκόλλου IEEE 802.1Q.

4.5 Μεθοδολογία Σχεδίασης

Στην πειραματική υποδομή του NOVI, αναπτύξαμε ένα μηχανισμό που μπορεί να ενώνει το FEDERICA με το PlanetLab/VINI. Οι αρχές σχεδίασης που ακολουθήσαμε και έχουν περιγραφεί νωρίτερα επιτρέπουν την ενοποίηση των εικονικοποιημένων δικτύων ενός ερευνητή/χρήστη και στις δυο υποδομές. Οι διαδικασίες απόδοσης και διαχείρισης των ετερογενών υποδομών έχουν υλοποιηθεί στο επίπεδο διαχείρισης του NOVI σε ένα αυτόνομο δομικό στοιχείο το οποίο ονομάζεται NSwitch Service. Το συγκεκριμένο δομικό στοιχείο είναι υπεύθυνο να επικοινωνεί με τις απαιτούμενες δικτυακές οντότητες των επιμέρους υποδομών δυναμικά. Μπορεί δηλαδή να δημιουργεί τις λογικές διεπαφές (logical interfaces) που απαιτούνται για την συνένωση

των εικονικών δικτύων των ερευνητών/χρηστών, όταν και εφόσον απαιτηθεί από τους ίδιους (ad-hoc).

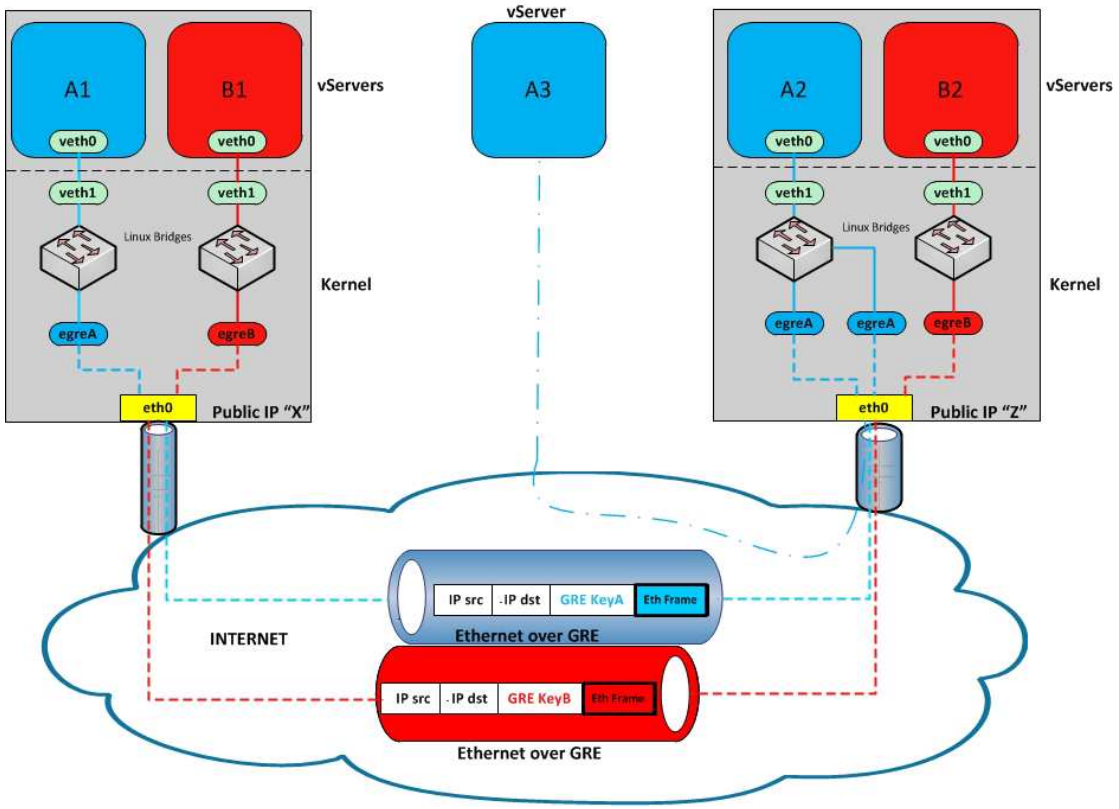
Στην πλευρά του PlanetLab/VINI, το οποίο χρησιμοποιείται ως παράδειγμα υποδομής που χρησιμοποιεί το Internet για να συνενώσει τις κατανεμημένες υποδομές του, χρησιμοποιούμε GRE tunnels για να δημιουργήσουμε εικονικές ζεύξεις μεταξύ των slivers. Η ενθυλάκωση Ethernet frames σε GRE tunnels (Ethernet over GRE) επελέγη λόγω της μικρής και σταθερής επιβάρυνσης που εισάγει και της ικανότητας ταυτόχρονης εξυπηρέτησης πολλαπλών ροών μέσα από την ίδια εικονική ζεύξη. Ένα συγκεκριμένο πεδίο της επικεφαλίδας του GRE πρωτοκόλλου, που ονομάζεται ***key filed*** [RFC2890] μπορεί να χρησιμοποιηθεί ως κριτήριο διαχωρισμού πολλαπλών ροών. Με αυτό τον τρόπο μετατρέπεται ένα GRE tunnel σε πολλαπλές εικονικές ζεύξεις, δυνάμενο να ικανοποιήσει πολλαπλούς ερευνητές/χρήστες.

Το *key field* είναι ένα πεδίο μήκους 32bit που συμπληρώνεται από τον encapsulator, το άκρο εισόδου των Ethernet frames δηλαδή μέσα στο GRE tunnel. **Το *key field* παρέχει το απαιτούμενο πλαίσιο για τον διαχωρισμό πολλαπλών ροών μεταξύ ενός συγκεκριμένου ζεύγους άκρων (endpoints), μετατρέποντας μια λογική ζεύξη (logical link) σε πολλαπλές εικονικές ζεύξεις (virtual links)**. Ethernet frames που ανήκουν στο ίδιο εικονικό δίκτυο (virtual network - slice) ενθυλακώνονται με χρήση της ίδιας ακριβώς τιμής στο πεδίο *key field,* με αποτέλεσμα να είναι δυνατή η ομαδοποίησή τους κατά την αποθυλάκωση (decapsulation) στο άκρο εξόδου των Ethernet frames σε μια ροή.

Η χρήση του *key field* έχει ιδιαίτερη σημασία όταν οι κόμβοι μιας υποδομής (π.χ PlanetLab) φιλοξενούν πολλαπλά virtual nodes και ο καθένας από αυτούς έχει μια δημόσια IP διεύθυνση (public IP), κοινόχρηστη για όλα τα virtual nodes. Σε αυτή την περίπτωση η υλοποίηση εικονικών ζεύξεων διαφορετικών slices μεταξύ ενός ζεύγους κόμβων απαιτεί τον διαχωρισμό της κίνησης σε ροές με χρήση του *key field.* Η κάθε ροή αντιστοιχίζεται με μια συγκεκριμένη εικονική ζεύξη που χαρακτηρίζεται από τον συνδυασμό/πλειάδα (tuple): *{source public IP, destination public IP, key field}*.

Ένα παράδειγμα διασύνδεσης μεταξύ virtual nodes που φιλοξενούνται σε ένα ζεύγος κόμβων και ανήκουν σε διαφορετικά εικονικά δίκτυα φαίνεται στο **Σχήμα 4**. Κάθε ένας από τους δυο φυσικούς κόμβους (physical nodes), σκιαγραφημένοι με γκρι χρώμα, φιλοξενεί δυο virtual nodes και ως λειτουργικό σύστημα χρησιμοποιεί το

Linux. Η ικανότητα του Linux bridge, δικτυακού στοιχείου του πυρήνα του λειτουργικού συστήματος, να δρα ως κλασσικός μεταγωγέας (MAC learning switch) με πολλαπλές διεπαφές επιτρέπει στον virtual node (A3) να είναι συνδεμένος στο ίδιο broadcast domain με τους A1 και A2, που συμμετέχουν στο ίδιο slice. Στο μέσο του **Σχήμα 4** απεικονίζονται δυο υποδείγματα IP πακέτων που ανήκουν σε δυο διαφορετικές ροές και κατ' επέκταση σε δυο διαφορετικά εικονικά δίκτυα. Αν και τα δυο άκρα του GRE tunnel, που μεταφέρει τα πακέτα, είναι ακριβώς τα ίδια τα πακέτα μπορούν να παραδοθούν στο σωστό Linux Bridge, εσωτερικά του πυρήνα του Λειτουργικού Συστήματος βάσει των τιμών στο πεδίο *key field (keyA* για το slice A, *keyB* για το slice B).



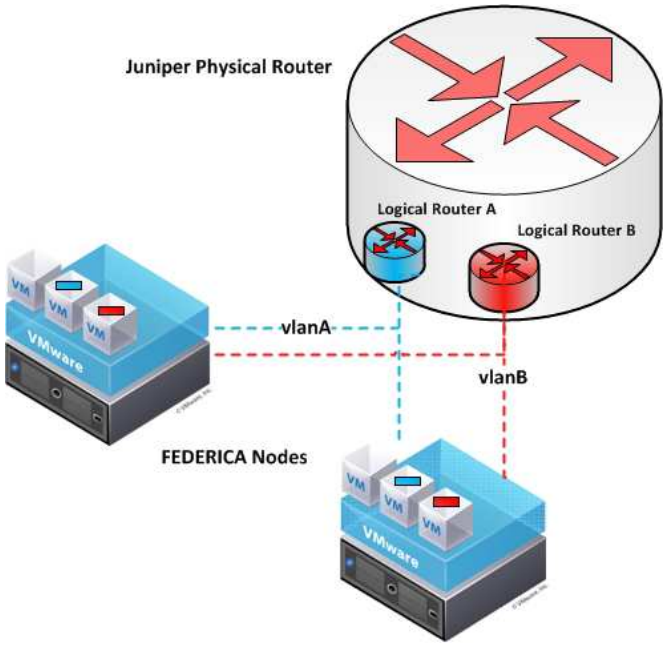
Σχήμα 4: Layer 2 διασύνδεση μεταξύ εικονικών κόμβων πάνω από το Internet, με χρήση Ethernet πλαισίων ελεγχόμενα από GRE tunnel και Linux switch (VINI/Trellis αρχιτεκτονική)

Με την συγκεκριμένη μεθοδολογία που ακολουθήσαμε επιτρέπουμε την ανάθεση μιας μοναδικής τιμής στο GRE *key field* ανά slice δημιουργώντας ένα μηχανισμό διαχωρισμού των virtual links στο επίπεδο φυσικής/λογικής υποδομής η οποία

χρησιμοποιείται για την εικονικοποίηση των Ethernet links, σεβόμενοι τις αρχές του Κεφαλαίου 3.1.3.

Όπως έχει ήδη αναφερθεί, υποδομές όπως το FEDERICA χρησιμοποιούν για την εικονικοποίηση των routers (logical routers) ειδικές εκδόσεις λειτουργικών συστημάτων που επιτρέπουν την ανεξάρτητη και παράλληλη λειτουργία πολλαπλών λογικών οντοτήτων που έχουν δικό τους επίπεδο ελέγχου, προώθησης δεδομένων και διαχείρισης. Στην περίπτωση του FEDERICA η διασύνδεση των φυσικών δρομολογητών με τους virtual nodes γίνεται από μεταγωγείς υλοποιημένους σε λογισμικό στο επίπεδο του λειτουργικού συστήματος VMware ESXi [vSwitch]. **Το πρωτόκολλο IEEE 802.1Q χρησιμοποιείται για να ενώσει τις δικτυακούς πόρους ενός slice, με το VLAN ID να είναι το κριτήριο διαχωρισμού μεταξύ logical links διαφορετικών slices**. Στο **Σχήμα 5** απεικονίζονται δυο slices, ένα με μπλε χρώμα και ένα με κόκκινο. Κάθε ένα περιλαμβάνει ένα logical router και δυο virtual nodes, τοποθετημένα σε δυο διαφορετικούς φυσικούς κόμβους. Το κάθε χρώμα αντιστοιχεί και σε μια μοναδική αποκλειστική τιμή VLAN ID που έχει συσχετισθεί με ένα και μόνο ένα slice.

Βασικό στοιχείο που χρησιμοποιήσαμε για την συνένωση του επιπέδου προώθησης δεδομένων είναι ο προγραμματιζόμενος Ethernet μεταγωγέας λογισμικού Open vSwitch (OVS). Οι επιδόσεις του OVS σε επίπεδο μεταγωγής πακέτων είναι εφάμιλλες με του Ethernet bridge του λειτουργικού συστήματος Linux [Pfaf09]. Η δυνατότητα δημιουργίας logical interfaces IEEE 802.1Q, IP over GRE, Ethernet over GRE, GRE over IPsec, η δυνατότητα χρήσης του OpenFlow στο επίπεδο ελέγχου και η ύπαρξη του Open vSwitch Database Management protocol (OVSDB) [OVSDB] στο επίπεδο διαχείρισης, το καθιστούν την πιο ολοκληρωμένη λύση μεταγωγέα λογισμικού που έχει αναπτυχθεί έως τώρα.



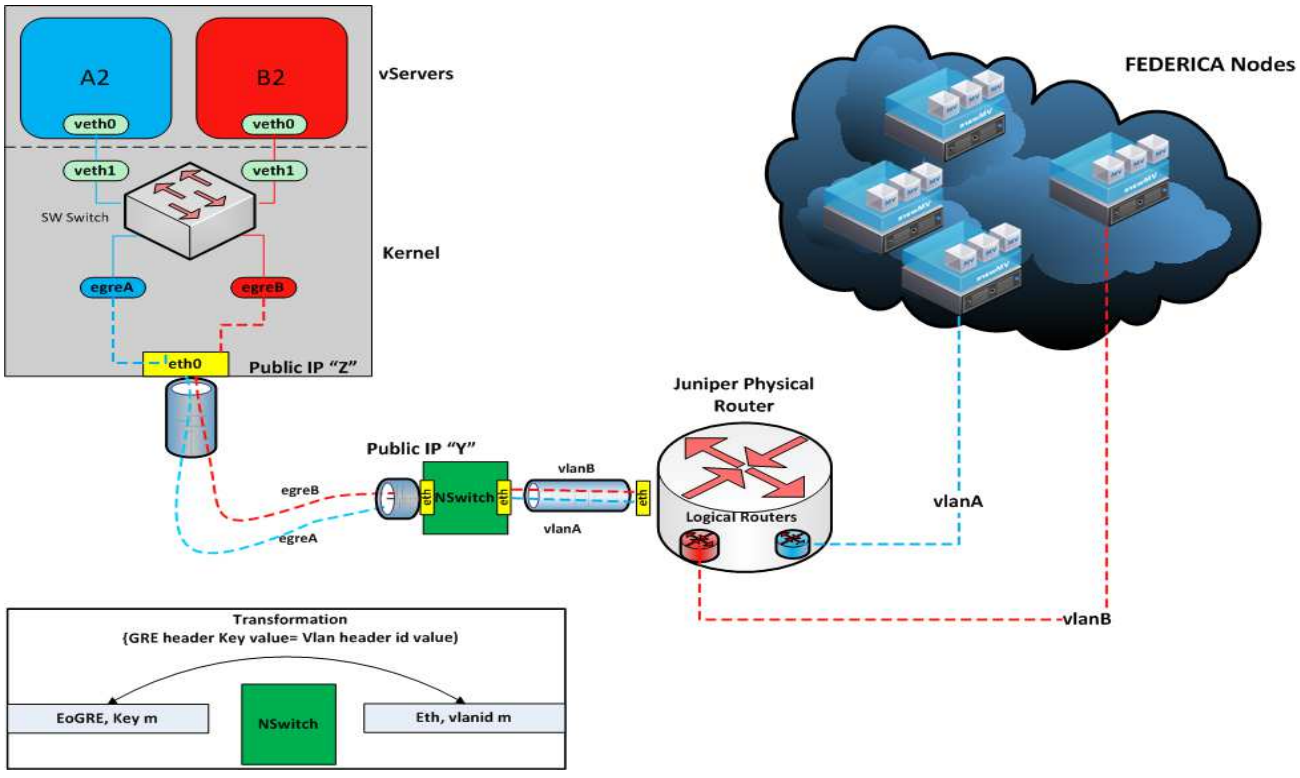
Σχήμα 5: Layer 2 διασύνδεση με χρήση VLANs μεταξύ εικονικών κόμβων τεχνολογίας VMware (πλήρης εικονικοποίηση) και λογικών δρομολογητών του Juniper MX-480

Για να γίνει η συνένωση του επιπέδου προώθησης δεδομένων χρησιμοποιήσαμε Ethernet over GRE logical interfaces και IEEE 802.1Q logical interfaces. Η τακτική που ακολουθήσαμε είναι η ομαδοποίηση υπό ένα broadcast domain των GRE logical interfaces που άνηκαν σε federated slice ερευνητή/χρήστη της μιας υποδομής και των IEEE 802.1Q logical interfaces που άνηκαν στο ίδιο federated slice της δεύτερης υποδομής. Η τεχνική υλοποίηση της συγκεκριμένης ομαδοποίησης έγινε με την εισαγωγή ενός VLAN ID (VID) σε κάθε GRE interface, ενοποιώντας ουσιαστικά σε Layer 2 τοπολογίες των δυο ετερογενών υποδομών και εξασφαλίζοντας την απομόνωση μεταξύ διαφορετικών slices.

Για να δημιουργήσουμε ένα περιβάλλον πολλαπλών χρηστών που θα υπάρχει απόλυτη αντιστοιχία μεταξύ slice και ενοποιημένης τοπολογίας χρειάστηκε να αναπτύξουμε μια τακτική για την αντιστοίχιση/ομαδοποίηση των ταυτοτήτων των virtual links στις ετερογενείς υποδομές. Στην υποδομή του PlanetLab χρησιμοποιήσαμε το GRE *key field* ως ταυτότητα (resource id) για την αντιστοίχιση slice και virtual link. Με αυτό τον τρόπο μετατρέψαμε το συγκεκριμένο πεδίο σε πεδίο δήλωσης ιδιοκτήτη – slice για κάθε virtual link. Το 32-bit εύρος του GRE *key field* επιτρέπει την χρήση του συγκεκριμένου πεδίου ως στοιχείο ταυτοποίησης ακόμα και σε πολύ μεγάλα ερευνητικά περιβάλλοντα (large-scale testbeds) που τα ταυτόχρονα πειράματα μπορεί να είναι χιλιάδες. Περιβάλλοντα όπως το FEDERICA, που χρησιμοποιούν ως στοιχείο

ταυτοποίησης το VLAN ID (VLAN network slicing) για τα virtual links, περιορίζονται από το εύρος των 12-bit, που μπορεί να διαφοροποιήσει σε 4095 (μια τιμή είναι δεσμευμένη) λογικές ομάδες τα Ethernet flows. Ο συγκεκριμένος αριθμός είναι επαρκής για περιβάλλοντα όπως το FEDERICA, αλλά όχι ικανός να εξυπηρετήσει ομοσπονδίες πολύ μεγάλης κλίμακας ή περιβάλλοντα μεγάλων data-centers και ειδικότερα υπηρεσίες Infrastructure as a Service (IaaS).

Για την άρση του συγκεκριμένου εγγενούς περιορισμού του VLAN network slicing σε ομόσπονδες πειραματικές υποδομές που χρησιμοποιούν διαφορετικές τεχνικές, δεσμεύσαμε τις τιμές από 1-4095 για συγκεκριμένη χρήση. Οι τιμές αυτές χρησιμοποιούνται ως ταυτότητα για slices που περιλαμβάνουν πόρους αποκλειστικά από υποδομές με VLAN slicing ή για slices που συνδυάζουν πόρους από υποδομές βασισμένες σε VLAN network slicing και GRE tunneling.



Σχήμα 6: Δικτυακή ομοσπονδοποίηση ετερογενών υποδομών σε Layer 2 με χρήση GRE & VLANs πρωτοκόλλων

Οι τιμές πάνω από το 4095 χρησιμοποιούνται ως ταυτότητα των virtual links ενός slice που περιλαμβάνει μόνο GRE tunnels ή άλλες τεχνολογίας δημιουργίας virtual links όπως για παράδειγμα το VXLAN [VXLAN], το οποίο όμως δεν έχει ακόμη γνωρίσει ευρεία χρήση.

Εκτός από την ταυτοποίηση και ομαδοποίηση των ετερογενών virtual links ανά slice θα πρέπει να υπάρχει και ένας μηχανισμός για την ταυτοποίηση της λογικής οντότητας (logical instance) πάνω στην οποία συνδέονται τα διαφορετικής τεχνολογίας virtual interfaces (Ethernet over GRE και VLAN logical interfaces).

Στην περίπτωση του OVS, το σύνολο των virtual links με ταυτότητες από 1-4095 μπορεί να συνδεθεί πάνω σε μια μοναδική λογική οντότητα, λόγω της εγγενούς δυνατότητας χαρακτηρισμού ενός GRE logical interface ως VLAN access port και της υποστήριξης του IEEE 802.1Q ως πρωτοκόλλου για τον διαχωρισμό των Ethernet frames. Στην περίπτωση που ένα GRE logical interface δεν μπορεί να χαρακτηρισθεί ως VLAN access port θα πρέπει να χρησιμοποιηθεί μια ξεχωριστή λογική οντότητα του OVS για κάθε νέο slice.

Για την σύνδεση virtual links που έχουν ταυτότητα με τιμή άνω του 4095 θα πρέπει να χρησιμοποιηθεί μια ξεχωριστή λογική οντότητα του OVS για κάθε νέο slice που περιλαμβάνει πόρους από δυο ή παραπάνω πειραματικές υποδομές.

Πλήθος δικτυακών κόμβων που υποστηρίζουν το πρωτόκολλο IEEE 802.1Q και την χρήση του VID, ως λογικού διαχωριστή διαφορετικών broadcast domains, δεν έχουν την ικανότητα δημιουργίας Ethernet over GRE logical interfaces. Κάποιες άλλες υποστηρίζουν Ethernet over GRE logical interfaces αλλά δεν έχουν την δήλωση ενός GRE logical interface ως VLAN access port και άρα την δημιουργία ενός broadcast domain μεταξύ ετερογενών logical interfaces. Η συγκεκριμένη παρατήρηση ισχύει και για την υποδομή του FEDERICA (MX-480 routers).

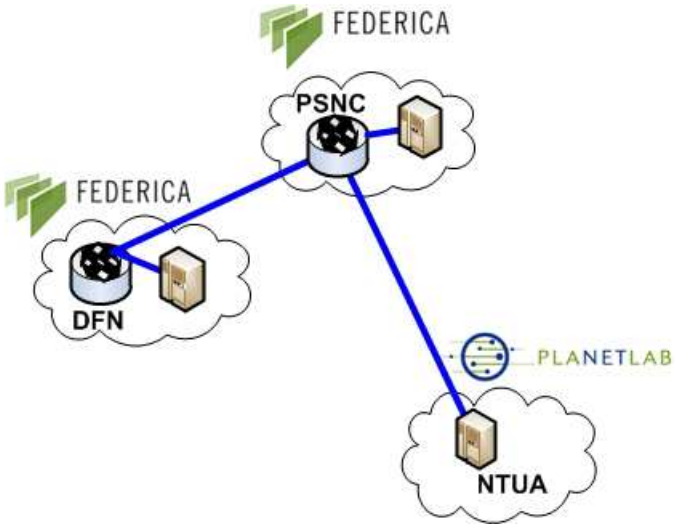
Οι αρχές, η μεθοδολογία και η υλοποίηση του προγραμματιζόμενου Ethernet μεταγωγέα λογισμικού υιοθετήθηκαν και δοκιμάσθηκαν στα πλαίσια του NOVI [Argy13]. Το όνομα του πρωτοτύπου που υλοποιήθηκε ονομάσθηκε NSwitch. Στο **Σχήμα 6** απεικονίζεται η συνολική εικόνα συνένωσης του επιπέδου προώθησης δεδομένων των δυο ετερογενών ομόσπονδων υποδομών δικτυακής εικονικοποίησης. Κάθε slice στην προκειμένη περίπτωση αποτελείται από slivers και από τις δυο

υποδομές. Συνδυάζοντας το μπλε slice του **Σχήμα 4** με αυτό του **Σχήμα 5** δημιουργούμε μια ενιαία τοπολογία, κρατώντας πλήρως απομονωμένα τα slices. Μπορούμε να παρατηρήσουμε ότι μέσα στο slice ανήκει και ένας logical router. Ανάλογα με την παραμετροποίηση του logical router οι virtual nodes μπορεί να ανήκουν στο ίδιο broadcast domain, δημιουργώντας στην απλούστερη περίπτωση ένα ενιαίο Local Area Network (Layer 2 forwarding) ή σε πολλαπλά broadcast domains με τον logical router να έχει λειτουργία ενός παραδοσιακού router (Layer 3 forwarding). Το ίδιο ισχύει και για το κόκκινο slice. Στο **Σχήμα 6** (κάτω αριστερά) διακρίνουμε επίσης την λογική αντιστοίχισης μεταξύ GRE interfaces και VLAN interfaces.

4.6 Πειραματικές Μετρήσεις

Σκοπός της εργασίας μας ήταν ο σχεδιασμός και η υλοποίηση μιας αποδοτικής λύσεως που δεν θα μείωνε τις επιδόσεις μιας δικτυακής τοπολογίας που θα εκτεινόταν σε μια ομοσπονδία υποδομών. Για την αξιολόγηση της επίδοσης της προτεινόμενης λύσης παρουσιάζουμε αποτελέσματα μετρήσεων που κάναμε στο πραγματικό περιβάλλον του NOVI. Το περιβάλλον που χρησιμοποιήσαμε για τις μετρήσεις περιλαμβάνει 3 σημεία παρουσίας της κατανεμημένης γεωγραφικά πειραματικής διάταξης. Το 1ο σημείο παρουσίας της πειραματική διάταξης ήταν στο Εθνικό Μετσόβιο Πολυτεχνείο (NTUA), το 2ο στο Poznan Supercomputing and Networking Center (PSNC) της Πολωνίας και το 3ο στο γερμανικό Δίκτυο Έρευνας και Τεχνολογίας (DFN).

Δημιουργήσαμε ένα πειραματικό slice με πόρους από τις υποδομές του FEDERICA και ενός ιδιωτικού αντίγραφου του PlanetLab που περιελάμβανε (i) ένα PlanetLab sliver στον NTUA (ii) ένα FEDERICA logical router στο PSNC (iii) ένα FEDERICA virtual node στο PSND (iv) ένα FEDERICA logical router στο DFN και (v) ένα virtual node στο DFN. Στη πλευρά του FEDERICA οι logical routers είναι υλοποιημένοι πάνω στους MX-480 routers και οι virtual nodes πάνω στην πλατφόρμα εικονικοποίησης ESXi. Η πειραματική τοπολογία φαίνεται στο **Σχήμα** 7. ανάμεσα στους φυσικούς πόρους οι οποίοι φιλοξενούσαν του συγκεκριμένους εικονικούς πόρους δίνοντάς μια άμεση εικόνα της αποδοτικότητας της λύσεως.



Σχήμα 7: Τοπολογία δοκιμών για εκτίμηση των επιδόσεων της λύσεως δικτυακής ομοσπονδοποίησης εικονικών κόμβων

Στα πειράματά μας συγκρίνουμε τους μέσους χρόνους μετάβασης και επιστροφής πακέτων (round-trip time –RTT) και το μέγιστο ρυθμό μετάδοσης που επιτεύχθηκε (maximum bandwidth). Τα αποτελέσματα των μετρήσεων ανάμεσα σε εικονικούς πόρους (sliver, logical router, virtual node) συγκρίνονται με τις αντίστοιχες τιμές που μετρήθηκαν

Στην 1η γραμμή του **Πίνακας 4-1** παρουσιάζονται οι τιμές που μετρήθηκαν για τους μέσους RTT χρόνους μεταξύ φυσικών πόρων και συγκεκριμένα του PlanetLab Host <NTUA PLHost>, του (i) FEDERICA PSNC Physical Router <PSNC PR>, (ii) FEDERICA PSNC ESXi Server <PSNC HV>, (iii) FEDERICA DFN Physical Router <DFN PR> και του (iv) FEDERICA PSNC ESXi Server <DFN HV>.

Στην 2η γραμμή (**Πίνακας 4-1**) παρουσιάζονται οι αντίστοιχοι χρόνοι μεταξύ των εικονικών πόρων που έχουν υλοποιηθεί πάνω από τους φυσικούς πόρους της 1η γραμμής (**Πίνακας 4-1**). Όπως παρατηρούμε υπάρχει μια μικρή υποβάθμιση της επιδόσεων στους RTT χρόνους που έχουν μετρηθεί στο επίπεδο της εικονικής τοπολογίας και των αντίστοιχων πόρων που είναι διασυνδεμένοι πάνω της. Για παράδειγμα ο μέσος RTT χρόνος μεταξύ NTUA PL Sliver> and <PSNC LR> αυξάνει από 83.4 msec (physical resources) σε 83.6 msec, δημιουργώντας μια επιβάρυνση της τάξεως του 0.2%. Η ίδια συμπεριφορά παρατηρείται για όλους τους RTT χρόνους στο επίπεδο του εικονικού δικτύου με την διακύμανση να κυμαίνεται μεταξύ 0.1% και 0.2%.

RTT (ms)	PSNC PR	PSNC LR	PSNC HV	PSNC VM	DFN PR	DFN LR	DFN HV	DFN VM
NTUA PlanetLab Host	83.4	-	83.8	-	116.7	-	117.1	-
NTUA PlanetLab Sliver	-	83.6 (-0.2%)	-	84.0 (-0.2%)	-	116.8 (-0.1%)	-	117.3 (-0.2%)

Πίνακας 4-1: Σύγκριση των μέσων χρόνων μετάβασης και επιστροφής πακέτων μεταξύ των φυσικών κόμβων και των αντίστοιχων εικονικών κόμβων τους

Ο **Πίνακας 4-2** αποτυπώνει τις τιμές μέγιστου ρυθμού μετάδοσης για TCP και UDP κίνηση μεταξύ φυσικών πόρων και της αντίστοιχης τοπολογίας τους καθώς και των εικονικών πόρων και της δικής τους υπερκείμενης εικονικής τοπολογίας. Όλες οι μετρήσεις που έγιναν είχαν ως πρώτο άκρο των δοκιμών τον <NTUA PLHost> και τον αντίστοιχο virtual node, <NTUA PLSliver>. Δεδομένου ότι δεν υπήρχε αξιόπιστος τρόπος να ληφθούν μετρήσεις στο επίπεδο των φυσικών πόρων της πλατφόρμας ESXi, με εγκατάσταση εργαλείων μετρήσεων ανοιχτού λογισμικού, λάβαμε μετρήσεις από το NSwitch στο PSNC για να τις συγκρίνουμε με αντίστοιχες τιμές από virtual nodes στο PSNC και στο DFN. Όπως παρατηρούμε το μέγιστο bandwidth μειώθηκε κατά 5,8 & και 5,6% για τον virtual node στο PSNC για TCP και UDP κίνηση αντίστοιχα. Λαμβάνοντας υπόψη ότι το μέγιστο bandwidth μειώνεται λόγω του γεγονότος ότι στις μετρήσεις εμπλέκεται και ένα virtual node, η μείωση που οφείλεται στην συνένωση των υποδομών με το NSwicth είναι κλάσμα των προαναφερθέντων ποσοστών. Τέλος, η μείωση του bandwidth του virtual node του DFN μπορεί να αποδοθεί στο γεγονός ότι συγκρίνεται με το bandwidth που καταφέραμε να επιτύχουμε χρησιμοποιώντας το NSwitch στο PSNC.

Bandwidth (Mbps)	NSwitch (PSNC)		PSNC VM		DFN VM	
	TCP	UDP	TCP	UDP	TCP	UDP
NTUA PlanetLab Host	81.5	95.1	-	-	-	-
NTUA PlanetLab Sliver	-	-	76.8 (-5.8%)	89.8 (-5.6%)	72.4 (-11.2%)	84.9 (-10.7%)

Πίνακας 4-2: Μέγιστος ρυθμός μετάδοσης μεταξύ των φυσικών κόμβων και των αντίστοιχων εικονικών κόμβων τους

5 Πλαίσιο Παθητικής Παρακολούθησης Υποδομών με Χρήση OpenFlow και sFlow στο Επίπεδο Διαχείρισης

Passive Flow Monitoring Framework for OpenFlow Enabled Experimental Facilities

5.1 Ερευνητικός Στόχος

Οι υποδομές που αναπτύσσονται τα τελευταία χρόνια προσπαθούν να ικανοποιήσουν τις ολοένα και αυξανόμενες απαιτήσεις των ερευνητών του Διαδικτύου του μέλλοντος. Για να καλυφθεί η ανάγκη ενός μηχανισμού απόδοσης, μέρους του επιπέδου ελέγχου της δικτυακής υποδομής στους ίδιους του ερευνητές/χρήστες, υιοθετήθηκε από ένα μέρος της ερευνητικής κοινότητας το πρωτόκολλο OpenFlow στο επίπεδο ελέγχου της δικτυακής υποδομή. Με αυτό τον τρόπο είναι δυνατός ο έλεγχος της προώθησης των πακέτων από ένα κεντρικοποιημένο επίπεδο ελέγχου, όπως έχει περιγραφεί στην Παράγραφο 2.4.

Η απαίτηση για την ταυτόχρονη χρήση της υποδομής από πολλαπλούς ερευνητές έκανε αναγκαία την ανάπτυξη ενός μηχανισμού τμηματοποίησης και απόδοσης του κεντρικοποιημένου ελέγχου σε πολλαπλές οντότητες, η κάθε μια ελεγχόμενη από διαφορετικό χρήστη. Ο μηχανισμός τεμαχισμού και απόδοσης μέρους του επιπέδου ελέγχου βασίστηκε πάνω στο μηχανισμό του ενδιάμεσου επιπέδου ελέγχου με χρήση ενός proxy OpenFlow controller (π.χ. FlowVisor).

Η παρακολούθηση και η συλλογή στατιστικών της δικτυακής κίνησης στις υποδομές που αναπτύχθηκαν [ProtoGeni] [GEANTOF] αρχικά αφέθηκε στις εγγενείς δυνατότητες του πρωτοκόλλου OpenFlow. Η συγκεκριμένη τακτική μεταφέρει την αρμοδιότητα παρακολούθησης της δικτυακής κίνησης από το επίπεδο διαχείρισης στο επίπεδο ελέγχου. Η ικανότητα κεντρικού ελέγχου ροών δεδομένων (data flows) με χρήση του OpenFlow επιτρέπει σε ένα μεγάλο εύρος εφαρμογών δικτύωσης, να προγραμματίζουν ουσιαστικά το δίκτυο. Για παράδειγμα μηχανισμοί firewalling, routing, load-balancing μπορεί να υλοποιηθούν κεντρικά στον OpenFlow controller

και να προγραμματίζουν τις απαραίτητες δικτυακές οντότητες κατευθύνοντας δυναμικά τις ροές δεδομένων.

Ερευνούμε κατά πόσο ένα πρωτόκολλο κεντρικοποιημένου ελέγχου όπως το OpenFlow μπορεί ταυτόχρονα να μεταφέρει όλες τις απαιτούμενες εντολές προώθησης ροών στις δικτυακές συσκευές και συγχρόνως να ικανοποιεί τις ανάγκες παρακολούθησης, όπως αρχικά είχε υποστηριχθεί από την ερευνητική κοινότητα.

Ελέγχουμε σε πειραματικό επίπεδο την αποτελεσματικότητα και την κλιμακοθετησιμότητα του OpenFlow στη λειτουργία της παθητικής παρακολούθησης του δικτύου για απαιτητικές εφαρμογές όπως είναι ο εντοπισμός δικτυακών ανωμαλιών (anomaly detection).Επίσης ελέγχουμε την ικανότητα αντιμετώπισης τέτοιων ανωμαλιών με χρήση του OpenFlow.

Τα συμπεράσματά μας είχαν ως αποτέλεσμα την ανάπτυξη ενός πλαισίου παθητικής παρακολούθησης υποδομών δικτυακής εικονικοποίησης ελεγχόμενες με το πρωτόκολλο OpenFlow που κάνει χρήση του sFlow και βοηθητικά άλλων πρωτοκόλλων όπως το SNMP. Το συγκεκριμένο πλαίσιο καλύπτει τις ανάγκες παρακολούθησης της δικτυακής κίνησης δίχως να δημιουργεί προβλήματα στο επίπεδο ελέγχου του δικτύου και στην ορθή λειτουργία των κεντρικοποιημένων λειτουργιών δικτύωσης.

Οι πειραματικές μετρήσεις που έγιναν σε πρωτότυπο μηχανισμού αντιμετώπισης δικτυακών ανωμαλιών δείχνει ότι η συλλειτουργία του sFlow και OpenFlow είναι δυνατή. Μηχανισμοί που κάνουν χρήση των δυο πρωτοκόλλων μπορεί να έχουν ευεργετικά αποτελέσματα όσον αφορά την κλιμακοθετησιμότητα απαιτητικών λειτουργιών, όπως αυτή του εντοπισμού δικτυακών ανωμαλιών, δίχως να μειώνεται αισθητά η αποτελεσματικότητα τους.

5.2 Αρχές Σχεδίασης

Το πλαίσιο παθητικής παρακολούθησης έχει ως στόχο την υποστήριξη της λειτουργίας ελέγχου του δικτύου όταν αυτή γίνεται με χρήση του πρωτοκόλλου OpenFlow. Πειραματικές υποδομές όπως το OFELIA [OFELIA] έχουν την δυνατότητα απόδοσης

του ελέγχου μιας ομάδας ροών μέσα στην φυσική τοπολογία τους στους ερευνητές/χρήστες. Ωστόσο δεν δίνουν την δυνατότητα παρακολούθησης των ροών παρά μόνο μέσα από τα στατιστικά των ροών που μπορεί να συλλέξει το OpenFlow.

Η παρακολούθηση των ροών είναι αναγκαία συνθήκη, για τον δυναμικό προγραμματισμό τους μέσα σε ένα κεντρικά ελεγχόμενο δίκτυο, καθώς ισχύει ο γενικός κανόνας ότι δεν μπορεί να βελτιστοποιήσεις δυναμικά κάτι που δεν μπορείς να το μετρήσεις.

Υπάρχουν δυο βασικές μέθοδοι παρακολούθησης, η πρώτη είναι η παθητική και η δεύτερη είναι η ενεργητική. Αν και οι δυο έχουν χρησιμοποιηθεί ευρέως σε δικτυακά περιβάλλοντα, θέλοντας να εξασφαλίσουμε την απομόνωση μεταξύ των διαφορετικών slices και να ελαχιστοποιήσουμε την δημιουργία δικτυακής κίνησης στο διαμοιραζόμενες φυσικές ζεύξεις, επιλέξαμε την παθητική προσέγγιση. Επιπλέον η παθητική παρακολούθηση ροών μπορεί να γίνει με χρήση πρωτοκόλλων όπως το SNMP [RFC3416], το NetFlow [RFC3954] και το sFlow [sFlow] που μπορεί να προσφέρουν στατιστικά στοιχεία ανά ροή, άμεσα αξιοποιήσιμα από έναν OpenFlow controller που έχει τον κεντρικό έλεγχου του δικτύου.

Βασιζόμενοι σε αυτό τον βασικό άξονα παρουσιάζουμε το Passive Flow Monitoring (PaFloMon) framework, με σκοπό τον εμπλουτισμό του επιπέδου διαχείρισης πειραματικών υποδομών Δικτυακής Εικονικοποίησης Ελεγχόμενες με το Πρωτόκολλο OpenFlow **(Σχήμα 9)**.

Οι βασικές αρχές σχεδίασης περιλαμβάνουν:

- Δυνατότητα παρακολούθησης ετερογενών δικτυακών υποδομών

  Το PaFloMon έχει την δυνατότητα παθητικής παρακολούθησης ασύρματων και ενσύρματων δικτυακών κόμβων.

- Δυνατότητα χρήσης πολλαπλών πρωτοκόλλων παρακολούθησης

  Η ανάγκη κάλυψης μεγάλους εύρους hardware συσκευών καθώς και η ανάγκη κάλυψης δικτυακών μεταγωγέων λογισμικού (π.χ OVS) που χρησιμοποιούνται κατά κόρον σε περιβάλλοντα εικονικοποίησης.

- Δυνατότητα αποθήκευσης, κατηγοριοποίησης και εμφάνισης των στατιστικών στοιχείων ανά slice

  Δυνατότητα συλλογής στατιστικών σε κατανεμημένα περιβάλλοντα, με τις λειτουργίες αποθήκευσης, κατηγοριοποίησης και εμφάνισης των στατιστικών στοιχείων ανά slice, ώστε να διατηρείται η ιδιωτικότητα (privacy) και η απομόνωση (isolation) ανά slice.

- Web-based και XML-RPC πρόσβαση των αποθηκευμένων στοιχείων

  Η δυνατότητα χρήσης ενός Web-based περιβάλλοντος απεικόνισης των στατιστικών ανά slice για πρόσβαση του ερευνητή και η παράλληλη υποστήριξη ενός αυτοματοποιημένου τρόπου πρόσβασης στα στατιστικά στοιχεία με χρήση XML Remote Procedure Calls (RPC).

- Ύπαρξη συγκεκριμένης προδιαγραφής παρακολούθησης για την ομογενοποίηση (homogeneity) και την εναρμόνιση (compatibility)

  Η υιοθέτηση μιας προδιαγραφής παρακολούθησης λύνει προβλήματα ομογενοποίησης της πληροφορίας στο πλαίσιο λειτουργίας του PaFloMon σε μια ομοσπονδία υποδομών, όπως αυτές που υποστηρίζει το OFELIA [OCF].

Η αποτελεσματικότητα ενός πλαισίου παθητικής παρακολούθησης υποδομών δικτυακής εικονικοποίησης με χρήση των πρωτοκόλλων OpenFlow και sFlow δοκιμάσθηκε με την ανάπτυξη πρωτότυπου μηχανισμού που σχεδιάσαμε για την αντιμετώπιση δικτυακών ανωμαλιών. Στόχος ήταν να ερευνήσουμε πόσο αποτελεσματικό είναι το συγκεκριμένο πλαίσιο για απαιτητικές εφαρμογές (π.χ. εντοπισμός δικτυακών ανωμαλιών) που ταυτόχρονα θα είχε ευεργετικά αποτελέσματα όπως:

- Βελτίωση της επεκτασιμότητας της λειτουργίας παθητικής παρακολούθησης, μέσω της χρήσης του sFlow, συγκριτικά με λύσεις που έχουν προταθεί και κάνουν χρήση του OpenFlow για ταυτόχρονο έλεγχο και παρακολούθηση του δικτύου.

- Αποτελεσματική μείωση της απαιτούμενης επικοινωνίας μεταξύ των OpenFlow switches και των OpenFlow controllers για τον περιορισμό του κινδύνου υπερφόρτωσης του επιπέδου ελέγχου, σε περιστατικά ανωμαλιών που μπορεί να παρατηρηθούν στο επίπεδο προώθησης δεδομένων από δικτυακές επιθέσεις ή φόρτο.

5.3 Ανασκόπηση της Υπάρχουσας Κατάστασης

Μέχρι στιγμής έχουν αναπτυχθεί αρκετά εργαλεία παρακολούθησης δικτυακών υποδομών και μερικά από αυτά έχουν χρησιμοποιηθεί σε πειραματικές διατάξεις για το Future Internet.

Η ομάδα του GENI με όνομα I&M (Instrumentation & Measurement) [GENIIM] έχει εργασθεί πάνω στο αντικείμενο της παρακολούθησης πειραματικών υποδομών. Μια από τις υπηρεσίες που αναπτύχθηκαν στα πλαίσια του GENI είναι το OnTimeMeasure [GENION] που δίνει την δυνατότητα οι ερευνητές/χρήστες να συντονίζουν και να συλλέγουν κεντρικοποιημένα αλλά και κατανεμημένα μετρήσεις που αφορούν το slice τους και περιλαμβάνουν παρακολούθηση των δικτυακών ζεύξεων, την εύρεση σημείων συμφόρησης (bottlenecks) καθώς και την διάγνωση σημείων που παρουσιάζονται λάθη. Ένα άλλο εργαλείο που έχει αναπτυχθεί κάτω από την ομπρέλα του GENI είναι το INSTOOL [GENIIN] που έχει την δυνατότητα παρακολούθησης παραμέτρων των πειραμάτων κατά την διάρκεια εκτέλεσής τους, με συλλογή στατιστικών με χρήση του SNMP. Τα βασικά στοιχεία που μπορεί να συλλέξει περιλαμβάνουν στοιχεία κίνησης (traffic statistics) και κόμβων (host statistics), την χρησιμοποίηση των ζεύξεων (network utilization).

Το GEMINI [GEMINI] παρέχει την δυνατότητα ενεργών και παθητικών μετρήσεων μέσα σε ένα GENI slice. Τα στατιστικά που συλλέγει περιλαμβάνουν στοιχεία μετρήσεων από κόμβους (CPU, memory) και εργαλεία για την μέτρηση της διεκπεραιωτικότητας (throughput), και της καθυστέρησης απλής μετάβασης (one-way delay) και μετάβασης μετ' επιστροφής πακέτων (round-trip time –RTT). Επίσης υποστηρίζει την μέτρηση του jitter και της απώλειας πακέτων (packet loss). Οι

ερευνητές/χρήστες μπορούν να προσδιορίσουν στο αίτημά τους ποιοι κόμβοι θα παρακολουθούνται μέσω παθητικής παρακολούθησης και μεταξύ ποίων θα γίνονται ενεργητικές μετρήσεις throughput και καθυστερήσεων.

Μια παρεμφερή προσπάθεια που έγινε για την υποστήριξη του ProtoGENI αποτελεί το Leveraging and Abstracting Measurements with perfSONAR (LAMP) [LAMP]. Το σύστημα μετρήσεων του LAMP βασίζεται στα εργαλεία του perfSONAR [perfSONAR] και εστιάζει το ενδιαφέρον του στην παροχή ενός κοινού και επεκτάσιμου format για την αποθήκευση δεδομένων και την ανταλλαγή τους. Το LAMP δεν μπορεί να ικανοποιήσει αιτήματα που προέρχονται από πολλαπλούς χρήστες.

Μια υπηρεσία παρακολούθησης που αναπτύχθηκε στα πλαίσια του GENI είναι το Scalable Sensing Service (S^3) [Blan12] το οποίο ακολουθεί μεθόδους ενεργητικής παρακολούθησης. Ο ερευνητής/χρήστης μπορεί να προγραμματίσει χρονικά την εκτέλεση των μετρήσεων ανά κόμβο που ανήκει σε ένα slice και να συλλέγει τα δεδομένα με αυτοματοποιημένο τρόπο.

Ένα σύνολο εργαλείων που αναπτύχθηκε και χρησιμοποιήθηκε στο PlanetLab είναι το Scriptroute suite [Spri02]. Ο στόχος της συγκεκριμένης ομάδας εργαλείων ήταν οι δικτυακές ρυθμίσεις και η εκσφαλμάτωση (debugging) μέσω μιας διεργασίας που έτρεχε στους PlanetLab nodes με αυξημένα δικαιώματα. Έδινε τη δυνατότητα στους χρήστες να διαχειρίζονται τα διακινούμενα πακέτα που τους άνηκαν στο δικό τους sliver.

Η πρώτη εξειδικευμένη υποδομή παρακολούθησης, για πειραματικές διατάξεις που κάνουν χρήση του πρωτοκόλλου OpenFlow στο επίπεδο ελέγχου, αναπτύχθηκε στο Πανεπιστήμιο Stanford [OFMF]. Δόθηκε βάρος στη συλλογή μετρήσεων σχετικών με το OpenFlow και με τις ροές που αυτό διαχειρίζεται, όπως ο χρόνος για την εγκαθίδρυση των flow entries στα OpenFlow switches, ο ρυθμός άφιξης ερωτημάτων για την εγκαθίδρυση νέων flow entries από τα OpenFlow switches στον OpenFlow controller, ενεργά flows ανά OpenFlow switch. Ήταν επίσης δυνατή η μέτρηση του RTT και καθυστερήσεων που εισήγαγε το κεντρικοποιημένο επίπεδο ελέγχου στην προώθηση των πακέτων. Για τις μετρήσεις και την συλλογή των στοιχείων γινόταν αποκλειστική χρήση κόμβων τοποθετημένων μέσα στο δίκτυο.

Οι μέχρι τώρα ερευνητικές προσπάθειες, για την ανάπτυξη μηχανισμών αντιμετώπισης δικτυακών ανωμαλιών σε δίκτυα οριζόμενα από λογισμικό, περιορίζονται σε μηχανισμούς που στηρίζονται μόνο στην χρήση του OpenFlow για παρακολούθηση. Η πρόταση που έγινε στο [Brag10] για τη δημιουργία μηχανισμού εντοπισμού ανωμαλιών σε δίκτυα οριζόμενα από λογισμικό βασίστηκε στην παρακολούθηση και έλεγχο ροών του δικτύου, αποκλειστικά με χρήση OpenFlow. Τα απαιτούμενα στατιστικά στοιχεία ανά flow συλλέγονταν από τους μετρητές που κρατά ένα OpenFlow switch στον OpenFlow controller που είχε στον έλεγχό του το δίκτυο. Η συγκεκριμένη εργασία δεν εξετάζει τυχόν περιορισμούς που εισάγονται από το μέγεθος του πίνακα ροών που έχουν στην διάθεσή τους τα OpenFlow switches, ούτε φαινόμενα υπερφόρτωσης του επιπέδου ελέγχου από την λειτουργία παρακολούθησης με την οποία έχει επιφορτιστεί το OpenFlow. Ειδικότερα, δεν γίνεται καμία αναφορά στο τρόπο που τυχόν επηρεάζουν ανωμαλίες στη δικτυακή κίνηση των χρηστών το κεντρικοποιημένο επίπεδο ελέγχου, το οποίο χειρίζεται ταυτόχρονα την παρακολούθηση της κίνησης αλλά και τις εντολές προώθησης δεδομένων με τις οποίες καθοδηγεί τις δικτυακές συσκευές.

Μια ακόμα προσπάθεια για τον εντοπισμό ανωμαλιών στο δίκτυο που κινείται στην κατεύθυνση παρακολούθησης και ελέγχου ροών του δικτύου μέσω του OpenFlow παρουσιάζεται στο [Mehd11]. Τα πειραματικά αποτελέσματα που παρουσιάζονται εστιάζουν στην αποτελεσματικότητα διαφόρων αλγορίθμων για την εύρεση των ανωμαλιών του δικτύου σε περιβάλλοντα χαμηλής κίνησης (οικιακά και εταιρικά περιβάλλοντα). Ο μέγιστος ρυθμός πακέτων που ο πειραματικός μηχανισμός του δοκιμάζεται φθάνει μέχρι τα 12.000 πακέτα. Στην περίπτωση του πειραματικού μηχανισμού που αναπτύξαμε για την δοκιμή της συνδυαστικής λειτουργία sFlow και OpenFlow ο ρυθμός πακέτων που χειριζόμαστε είναι της τάξεως των 120.000. Η συγκεκριμένη τάξη μεγέθους μπορεί να ικανοποιήσει δίκτυα με μεγαλύτερα φορτία κίνησης (π.χ. πανεπιστημιακά δίκτυα και μεγάλα εταιρικά δίκτυα). Η ανάλυσή μας καλύπτει την διερεύνηση των περιορισμών που εισέρχονται λόγω μεγέθους των flow tables και των κινδύνων υπερφόρτωσης του επιπέδου ελέγχου από πολλαπλά μηνύματα ελέγχου και στατιστικών στοιχείων (π.χ τιμές μετρητών) που ανταλλάσσονται σε τέτοιες συνθήκες λειτουργίας.

5.4 Το Ερευνητικό Περιβάλλον του OFELIA

Οι περισσότερες πειραματικές υποδομές που έχουν αναπτυχθεί στα πλαίσια του Future Internet υλοποιούν ένα σύνολο λειτουργιών του επιπέδου διαχείρισης και ελέγχου ώστε να είναι αφενός δυνατή η διάθεση των πόρων στους ερευνητές/χρήστες και αφετέρου να παρέχουν ένα σύνολο υπηρεσιών προς τους ερευνητές.

Με οδηγό την προαναφερθείσα λογική έχει αναπτυχθεί και το OFELIA control framework (OCF) που χρησιμοποιήσαμε ως παράδειγμα υποδομής για το πλαίσιο παθητικής παρακολούθησης PaFloMon. Το OFELIA προσφέρει τις υπηρεσίες του ταυτόχρονα σε πολλαπλούς χρήστες με χρήση τεχνολογιών εικονικοποίησης και slicing του επιπέδου ελέγχου μέσω του OpenFlow και του FlowVisor. Έχει την δυνατότητα πιστοποίησης (authentication) των χρηστών του και απόδοσης διαφορετικών ρόλων σε αυτούς (role-based authorization). Η υποδομή του υποστηρίζεται από δικτυακές υπηρεσίες όπως είναι το DNS και το VPN service για τους χρήστες του. Η γραφική διεπαφή που προσφέρεται στον χρήστη διευκολύνει την διαχείριση των virtual nodes καθώς και την ανάθεση σε συγκεκριμένο OpenFlow controller που ανήκει στον ερευνητή/χρήστη του flowspace. Το flowspace αποτελεί το μέρος εκείνο του επιπέδου ελέγχου που αποδίδεται σε ένα και μόνο slice συγκεκριμένου χρήστη με την βοήθεια του FlowVisor.

5.5 Μεθοδολογία Σχεδίασης

5.5.1 Πλαίσιο Παθητικής Παρακολούθησης

Το PaFloMon προβλέπει ως μηχανισμό πιστοποίησης και εξουσιοδότησης το Light-weight Directory Access Protocol (LDAP) [Kout04]. Η λειτουργία του απαιτεί την τροποποίηση του LDAP schema της εκάστοτε υποδομής με ένα σύνολο χαρακτηριστικών/παραμέτρων (attributes) που σχετίζονται με την λειτουργία της παθητικής παρακολούθησης. Μπορεί μέσω των συγκεκριμένων παραμέτρων να ελέγχει ποιοι ερευνητές/χρήστες έχουν δικαίωμα να χρησιμοποιούν συγκεκριμένα

πρωτόκολλα παρακολούθησης, όπως για παράδειγμα το sFlow, τις αντίστοιχες παραμέτρους τους (π.χ ρυθμός δειγματοληψίας). Μπορεί επίσης να κάνει διαχωρισμό στα δικαιώματα μεταξύ ερευνητών/χρηστών του ίδιου slice που διαχειρίζονται τα δεδομένα του slice τους.

Η σχεδίαση του PaFloMon προβλέπει τα κάτωθι:

- Συλλογή δεδομένων παθητικής παρακολούθησης (passive monitoring data collection)

  Μπορεί να συλλέγει δεδομένα παθητικής παρακολούθησης της δικτυακής κίνησης από πρωτόκολλα όπως είναι το sFlow και το NetFlow καθώς και δεδομένων που σχετίζονται με την κατάσταση των OpenFlow switches με χρήση του SNMP.

- Διαχείριση των agents παθητικής παρακολούθησης των OpenFlow switches

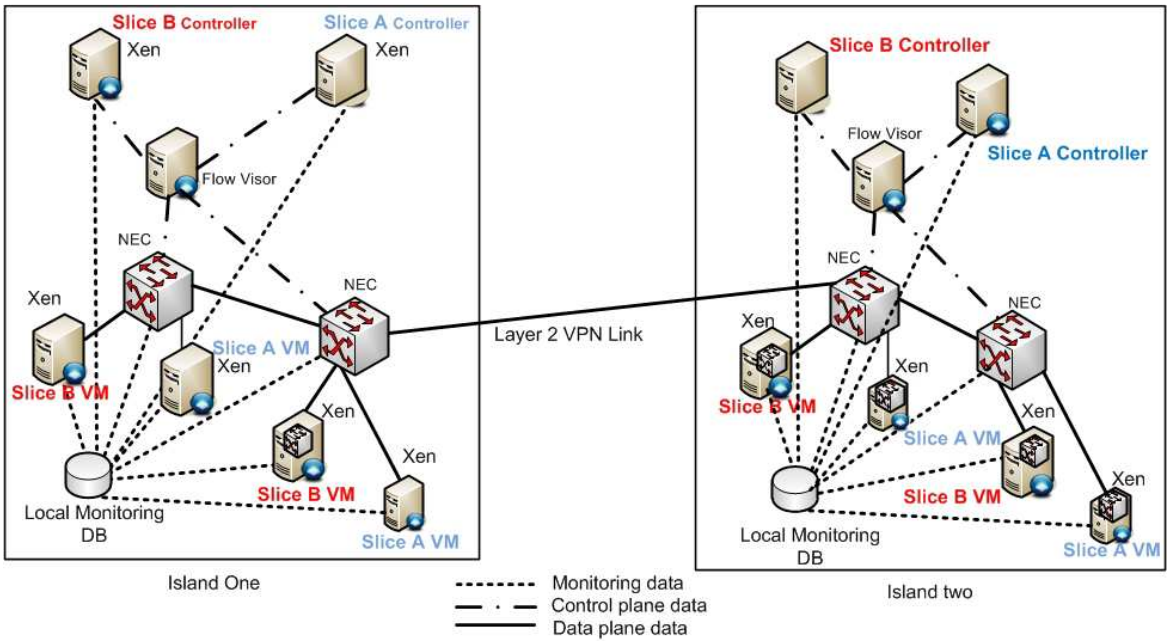
  Η χρήση των sFlow, NetFlow agents ανά OpenFlow switch και η δυναμική τους παραμετροποίηση (π.χ. ρυθμός δειγματοληψίας).

- Διαχωρισμό των δεδομένων ανά slice (data isolation per slice)

  Τα δεδομένα παρακολούθησης της δικτυακής κίνησης και των OpenFlow switches οργανώνονται μετά την συλλογή τους με τέτοιο τρόπο ώστε να μπορεί να γίνει ο διαχωρισμός τους ανά slice. Κάθε ερευνητής/χρήστης έχει πρόσβαση μόνο στα δεδομένα που αφορούν το slice του. Τα δεδομένα ομαδοποιούνται με κριτήρια τον τύπο του agent συλλογής δεδομένων και τα switches από τα οποία έχει γίνει η συλλογή.

- Δυνατότητα συλλογής δεδομένων παθητικής παρακολούθησης κόμβων του ερευνητή

  Ο ερευνητής/χρήστης μπορεί να παραμετροποιήσει τυχόν δικτυακούς κόμβους που έχει προσθέσει στο slice του (π.χ με χρήση VPN tunnels) με τέτοιο τρόπο ώστε να αποστέλλουν τα δεδομένα τους στον server συλλογής του PaFloMon. Σε αυτή την περίπτωση δεν μπορεί να γίνει δυναμική παραμετροποίηση των πρόσθετων δικτυακών κόμβων για παραμέτρους όπως είναι το packet sampling από το PaFloMon.



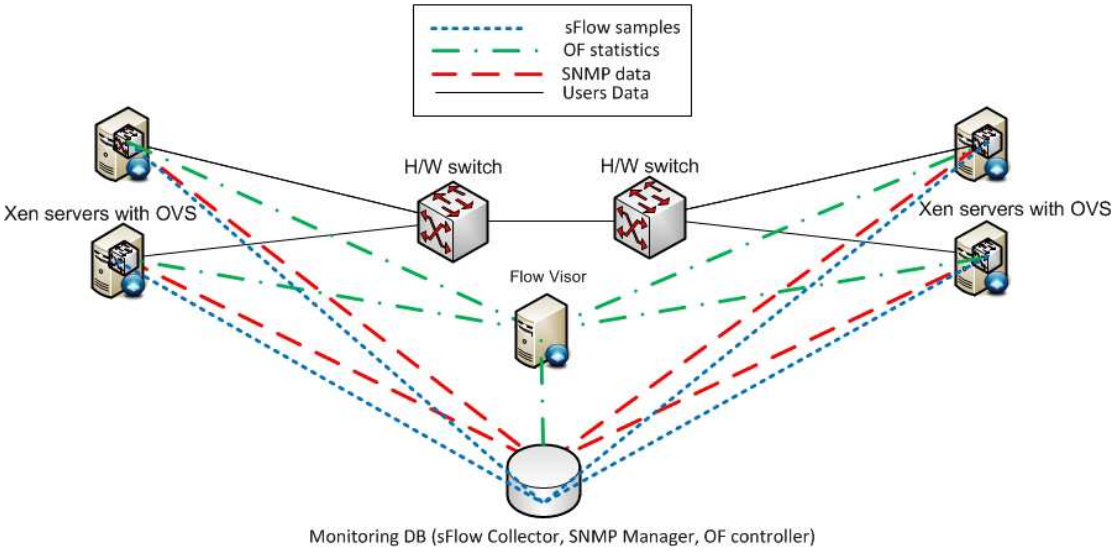
Σχήμα 8: Συλλογή δεδομένων παθητικής παρακολούθησης με χρήση sFlow collector, OpenFlow controller και SNMP manager σε κεντρική βάση δεδομένων

Τα OpenFlow switches σε ένα περιβάλλον δικτυακής εικονικοποίησης μπορεί να είναι υλοποιημένα είτε σε hardware είτε σε software που τρέχει πάνω σε υπολογιστές γενικού σκοπού. Στην πρώτη κατηγορία ανήκουν όλες εκείνες οι hardware συσκευές που κάνουν την υλοποίηση με χρήση κυκλωμάτων ειδικού σκοπού (network processors, μνήμες ternary content-addressable memory) και στην δεύτερη ανήκουν

οι υλοποιήσεις των software switches για λειτουργικά συστήματα που υποστηρίζουν εικονικοποίηση (hypervisors' operating systems).

Πολλά από τα OpenFlow hardware switches υποστηρίζουν πρωτόκολλα όπως είναι το SNMP και το sFlow, με χρήση των οποίων μπορεί να γίνει η παρακολούθηση του δικτύου. Το sFlow περιλαμβάνει ένα διαφορετικό μηχανισμό σταθμοσκόπησης των μετρητών (counter polling). Ο sFlow agent μπορεί περιοδικά να αποστέλλει τις τιμές των μετρητών σε έναν εξωτερικό συλλέκτη (collector). Η περιοδική αποστολή των τιμών είναι αποδοτικότερη από την συλλογή των αντίστοιχων δεδομένων με SNMP, αφού χρειάζεται πολύ μικρότερος αριθμός πακέτων (10 έως 20 φορές) για την συλλογή των ίδιων δεδομένων. Ο βασικός λόγος που υπάρχει αυτή η διαφορά στην απόδοση είναι ότι το sFlow χρησιμοποιεί το XDR [RFC4506] πρότυπο για την κωδικοποίηση (encode) των τιμών το οποίο είναι απλούστερο αυτού που προσφέρεται

από το ASN.1 [RFC3641], με αποτέλεσμα να μειώνεται ο φόρτος επεξεργασίας στο switch και στον συλλέκτη. Επίσης ένα ακόμα χαρακτηριστικό είναι ότι το sFlow παρέχει δείγματα πακέτων (packet samples), με ρυθμό που καθορίζεται ως παράμετρος κατά την λειτουργία του πρωτοκόλλου και είναι παραμετροποιήσιμος.



Σχήμα 9: Το επίπεδο παθητικής παρακολούθησης για πειραματικές υποδομές δικτυακής εικονικοποίησης ελεγχόμενες με το πρωτόκολλο OpenFlow

Σε μεγάλα μάλιστα δίκτυα, που ο ρυθμός μετάδοσης πακέτων είναι πολύ μεγάλος (πάνω από 10 Gbit/sec) η παραμετροποίηση του ρυθμού αποστολής δειγμάτων πακέτων γίνεται ακόμη χρησιμότερο χαρακτηριστικό.

Η ενεργοποίηση των sFlow agents σε επίπεδο OpenFlow switches της πειραματικής υποδομής περιλαμβάνει και την παραμετροποίηση του ρυθμού αποστολής δειγμάτων στον συλλέκτη. Στις πειραματικές μας μετρήσεις περιλαμβάνουμε μετρήσεις που έγιναν για την παραμετροποίηση της συγκεκριμένης παραμέτρου στα OpenFlow switches του OFELIA (NEC IP8800). Τονίζεται ότι η συγκεκριμένη παράμετρος αφορά το switch πάνω στο οποίο έχει υλοποιηθεί το sFlow (switch-wide) και δεν μπορεί να γίνει από το ίδιο το switch διαχωρισμός των συλλεχθέντων δεδομένων ανά slice.

Το sFlow παράλληλα με την χρήση του OpenFlow **(Σχήμα 9)** ως πρωτοκόλλου ελέγχου μπορεί να λειτουργεί επικουρικά παρέχοντας στατιστικά για την κίνηση του δικτύου. Τα στατιστικά που μπορεί να εξάγονται ανά τακτά χρονικά διαστήματα

ωφελούν πειράματα πάνω στον έλεγχο δικτύου. Για παράδειγμα αλγόριθμοι data-fusion για τον εντοπισμό και εξομάλυνση επιθέσεων λαμβάνουν υπόψη τους δεδομένα παθητικής παρακολούθησης από το δίκτυο [Siat05] [Andr09].

Η γενικότερη συνδρομή των δεδομένων παθητικής παρακολούθησης σε ένα δίκτυο ελεγχόμενο από το OpenFlow είναι ότι παρέχεται ανάδραση (feedback) για την κατάσταση του δικτύου, η οποία σε συνδυασμό με την δυναμική προώθηση ροών (dynamic flow forwarding) που προσφέρει το OpenFlow κάνουν το δίκτυο προγραμματιζόμενο (programmable).

Η δικτυακή διασύνδεση των virtual machines απαιτεί χρήση software switches μέσα στους hypervisors που φιλοξενούν τα virtual machines. Γνωστές υλοποιήσεις είναι το NEXUS 1000V της Cisco [NEXUS] και το Open vSwitch που ήδη έχει αναφερθεί στην Παράγραφο 4.2. Το OFELIA επέλεξε να κάνει χρήση του Linux bridge στα πρώτα στάδια υλοποίησής του και μετέπειτα χρησιμοποίηση το OVS στου XEN hypervisors [XEN].

Η μέθοδος παθητικής παρακολούθησης που προτείνουμε μπορεί να εκμεταλλευτεί τα πρωτόκολλα παθητικής παρακολούθησης (sFlow, NetFlow, SNMP) πέραν των στατιστικών που το OpenFlow προσφέρει για τις ροές δεδομένων που έχουν εγκαθιδρυθεί μέσα στα OpenFlow switches. Συγχρόνως παρακολουθείται η επιβάρυνση που προκύπτει στα OpenFlow switches από την χρήση των συγκεκριμένων πρωτοκόλλων για να μην γίνεται υπερφόρτωση της λειτουργίας τους, ειδικά στην περίπτωση των software switches που είναι υλοποιημένα μέσα στους hypervisors και θα μπορούσε να έχει ως αποτέλεσμα την δυσλειτουργία των εικονικών μηχανών ή της δικτυακής διασύνδεσής του με την υπόλοιπη τοπολογία.

Επιπλέον, με την χρήση των Host sFlow [sFlowHost] στατιστικών είναι δυνατή η παρακολούθηση των hypervisors και των virtual machines. Οι Host sFlow δομές επιτρέπουν την παρακολούθηση χρήσης του CPU, της μνήμης και των αποθηκευτικών μέσων. Ο συνδυασμός των στατιστικών δεδομένων από τους hypervisors και τα virtual resources (compute resources) με τα στατιστικά από τα OpenFlow switches (network resources) μπορούν να δώσουν μια συνολική εικόνα λειτουργίας των υπηρεσιών ή/και των πειραμάτων που εξυπηρετούνται από την υποδομή, κάνοντας αποτελεσματικότερη την αξιοποίηση των πόρων (π.χ load-balancing).

Μια άλλη κατηγορία στατιστικών στοιχείων που συλλέγονται με την χρήση του sFlow είναι αυτά που αφορούν ασύρματους κόμβους και περιλαμβάνουν αριθμό συγκρούσεων (collisions) ανά ασύρματη διεπαφή, λόγο σήματος προς θόρυβο (signal-to-noise ratio) κλπ. Σε πειραματικές υποδομές που γίνεται η χρήση ασύρματων κόμβων και του OpenFlow για τον έλεγχο ροών, μπορεί να προσφέρει στατιστικά για την δυναμική προώθηση των ροών στο ασύρματο δίκτυο που σχηματίζουν οι κόμβοι [sFlow802.11].

Ο σχεδιασμός του PaFloMon περιλαμβάνει την δημιουργία ενός κεντρικού χώρου αποθήκευσης (data repository) των δεδομένων παρακολούθησης που συλλέγονται για τα διάφορα slices. Το repository διατηρεί τα δεδομένα, διαχωρίζοντας αυτά με λογικό τρόπο ώστε να είναι δυνατή η πρόσβαση σε αυτά ανά slice. Δημιουργείται δηλαδή μια κατηγοριοποίηση ανά slice βάσει του slice id (π.χ VLAN id ή GRE key field). Επίσης υπάρχει περεταίρω ομαδοποίηση αναλόγως της μεθόδου συλλογής και των μετρικών που χρησιμοποιήθηκαν.

Για την δημιουργία ενός σχήματος περιγραφής (monitoring schema) επιλέξαμε να χρησιμοποιήσουμε το Resource Specification schema (RSpec) [RSpec] που κάνει χρήση της Extensible Markup Language (XML). Το RSpec χρησιμοποιείται για την περιγραφή από πλευράς χρήστη των δεδομένων παθητικής παρακολούθησης στα οποία θέλει να έχει πρόσβαση (αριθμός πακέτων, δείγματα πακέτων, στατιστικά ροών). Τα δεδομένα συλλέγονται και γίνεται η διαλογή τους ώστε τελικά να αποθηκεύονται μόνο τα δεδομένα που έχουν ζητηθεί από τους χρήστες. Η συγκεκριμένη διαλογή είναι απαραίτητη λόγω του τρόπου λειτουργίας των collectors (π.χ sFlow), δεδομένου ότι συλλέγονται δεδομένα τα οποία μπορεί να μην έχουν ζητηθεί (push model) και άρα να μην είναι απαραίτητο να αποθηκευτούν καταναλώνοντας αποθηκευτικό χώρο.

5.5.2 Μηχανισμός για την Αντιμετώπιση Δικτυακών Ανωμαλιών

Η ανάπτυξη πρωτότυπου μηχανισμού του πλαισίου παθητικής παρακολούθησης υποδομών δικτυακής εικονικοποίησης με χρήση των πρωτοκόλλων OpenFlow και

sFlow για την αντιμετώπιση δικτυακών ανωμαλιών/επιθέσεων βασίστηκε στις παρακάτω σχεδιαστικές αρχές:

- Αρθρωτή σχεδίαση με συλλογή των δεδομένων, εντοπισμό των δικτυακών ανωμαλιών και εξομάλυνση τους να εκτελούνται ανεξάρτητα

  Οι τρεις (3) λειτουργίες παρουσιάζονται στο **Σχήμα 10.** Αρχικά γίνεται η συλλογή των στατιστικών στοιχείων για τις ροές του δικτύου, ακολουθεί ο εντοπισμός των ανωμαλιών της κίνησης βάσει κάποιου αλγορίθμου και τέλος ακολουθεί η εξομάλυνση της ανωμαλίας με χρήση του πρωτοκόλλου OpenFlow και παύση των ροών που έχουν δημιουργηθεί από την επίθεση.

  Ο συλλέκτης των δεδομένων περιοδικά εξάγει τα στοιχεία των ροών στο επόμενο δομικό στοιχείο της επεξεργασίας. Η σχεδίαση περιλαμβάνει 2 διαφορετικούς συλλέκτες, εκ των οποίων ο πρώτος συλλέγει περιοδικά OpenFlow δεδομένα από τα flow tables και ο δεύτερος χρησιμοποιεί δείγματα πακέτων που αποστέλλονται από το switch με χρήση του sFlow. Στις πειραματικές μετρήσεις χρησιμοποιήθηκαν και οι 2 συλλέκτες για εξαγωγή συγκριτικών στοιχείων και έλεγχο της αποτελεσματικότητάς τους.

  Τα δεδομένα ροών που προωθούνται στο δομικό στοιχείο της επεξεργασίας ελέγχονται κατά διαστήματα (παράθυρο δευτερολέπτων) για δικτυακές ανωμαλίες οι οποίες φανερώνουν κίνηση προερχόμενη από επίθεση ή προετοιμασία για επίθεση (π.χ. διαδικασία σάρωσης πορτών – port scanning). Αλγόριθμοι εντοπισμού στατιστικών ανωμαλιών [Stan02], machine learning αλγόριθμοι [Ahme07] και data mining αλγόριθμοι [Wu09][Teod09][Patc07] μπορούν να χρησιμοποιηθούν για την εξεύρεση των δικτυακών ανωμαλιών, όπως αυτοί αναφέρονται στην βιβλιογραφία. Για τις πειραματικές μετρήσεις που έγιναν στο πρωτότυπο μας, χρησιμοποιήθηκε αλγόριθμος εντοπισμού ανωμαλιών βασιζόμενος στην εντροπία των τιμών των επικεφαλίδων των πακέτων [Lakh05] [Andr09]. Με την χρήση ενός τέτοιου αλγορίθμου είναι δυνατός ο εντοπισμός διαφορετικών ειδών επιθέσεων (π.χ port scan, distributed denial of service – DDoS, worm propagation, port scans).

Μετά τον εντοπισμό του κόμβου, που δημιουργεί την δικτυακή ανωμαλία από δομικό στοιχείο της επεξεργασίας γίνεται ενημέρωση του δομικού στοιχείου εξομάλυνσης. Σε αυτό το στάδιο δημιουργούνται οι κατάλληλοι κανόνες ροών (flow entries) που έχουν στόχο την απόρριψη των πακέτων που εμπίπτουν σε αυτές τις ροές. Η ενημέρωση των δικτυακών συσκευών με αυτούς του νέους κανόνες υψηλής προτεραιότητας γίνεται με την χρήση του OpenFlow. Από την στιγμή εγκαθίδρυσης των κανόνων και επειδή οι συγκεκριμένοι κανόνες έχουν υψηλότερη προτεραιότητα από κάθε άλλο κανόνα προώθησης πακέτων, η κακόβουλη ροή πακέτων σταματά.

- Συμβατότητα με οποιαδήποτε Layer 2 ή Layer 3 συσκευή

  Η αρθρωτή αρχιτεκτονική κρατά την συλλογή των δεδομένων ανεξάρτητη από τον εντοπισμό των ανωμαλιών. Για τη συλλογή στατιστικών των ροών μπορεί να χρησιμοποιηθεί οποιαδήποτε συσκευή υποστηρίζει πρωτόκολλα παθητικής παρακολούθησης όπως το sFlow και το NetFlow [NetFlow].

- Απαλοιφή των περιορισμών που εισάγει η χρήση στατιστικών ροών που προέρχονται αμιγώς από το OpenFlow

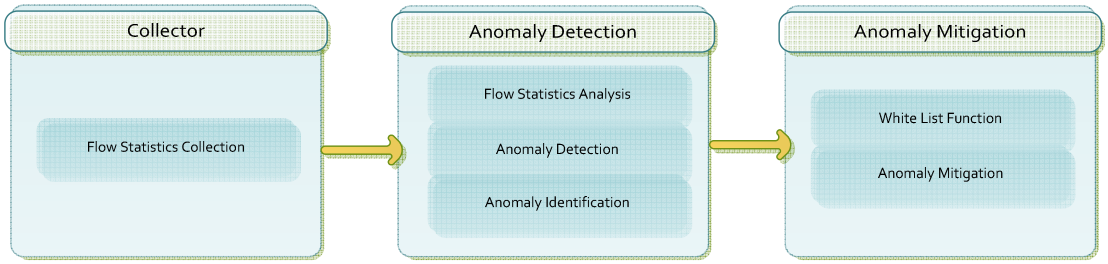
  Σε δίκτυα οριζόμενα από λογισμικό, όπου η συλλογή στατιστικών γίνεται μέσω του OpenFlow, απαιτείται συλλογή στοιχείων από τους counters των ροών. Οι counters κρατούν στοιχεία για κάθε flow entry που χρησιμοποιείται για την προώθηση δεδομένων, με αποτέλεσμα να είναι δυνατή καταγραφή στατιστικών μόνο για ροές που είναι αποθηκευμένες μέσα στα flow tables. Ο περιορισμένος αριθμός εγγραφών στο flow table, λόγω δυσκολιών στην οικονομική υλοποίηση των flow tables σε hardware, δημιουργεί την ανάγκη για την συνάθροιση ροών (aggregation) και των αντίστοιχων εντολών προώθησης δεδομένων. Η συνάθροιση ροών σε επίπεδο προώθησης είναι αποδεκτή και γίνεται κατά κόρον (π.χ routing tables) σε υπάρχουσες συσκευές. Η απαίτηση όμως για εντοπισμό ροών που έχουν ως σκοπό δικτυακές επιθέσεις απαιτεί την καταγραφή ροών που ορίζονται και από πεδία επικεφαλίδων στο Layer 4. Ο ορισμός των ροών με επικεφαλίδες του Layer 4 (microflows) εκτινάσσει τον συνολικό τους αριθμό ανά συσκευή και δημιουργείται πρόβλημα λόγω του αριθμού τους καθώς δεν μπορούν να φιλοξενηθούν στο flow table.

- Εντοπισμός των δικτυακών ανωμαλιών και εξομάλυνσή τους σε πραγματικό χρόνο λόγω αποσύζευξης του επιπέδου προώθησης δεδομένων και του επιπέδου ελέγχου

  Ο εντοπισμός δικτυακών ανωμαλιών απαιτεί μεγάλη επεξεργαστική ισχύ και δεν είναι δυνατόν να εκτελείται πάνω στις δικτυακές συσκευές. Συγχρόνως η εξομάλυνση των ανωμαλιών σε πραγματικό χρόνο απαιτεί έναν μηχανισμό που μπορεί να ελέγχει την λογική προώθησης των πακέτων. Ένας OpenFlow controller μπορεί να εισάγει κανόνες απόρριψης συγκεκριμένων ροών όταν αυτές έχουν χαρακτηριστεί ως επιβλαβείς από τον αλγόριθμο εντοπισμού ανωμαλιών.

- Χρήση τεχνικών δειγματοληψίας για αύξηση της κλιμακοθετησιμότητας του μηχανισμού εντοπισμού και εξουδετέρωσης επιθέσεων.

  Η σχεδίασή μας επιτρέπει την αποσυμφόρηση των δικτυακών συσκευών από λειτουργίες παθητικής παρακολούθησης λόγω της δειγματοληψίας που επιτελείται, με αποτέλεσμα να γίνεται δυνατή η παρακολούθηση μεγάλου όγκου κίνησης χωρίς τη χρήση συσκευών ειδικού σκοπού.



Σχήμα 10: Αρθρωτή σχεδίαση μηχανισμού συλλογής των δεδομένων, εντοπισμού των δικτυακών ανωμαλιών και εξομάλυνση τους σε δίκτυα οριζόμενα από λογισμικό

5.6 Πειραματικές Μετρήσεις

5.6.1 Πειραματικές Μετρήσεις Παθητικής Παρακολούθησης

Για να μελετήσουμε το κατά πόσο είναι εφικτή η παθητική παρακολούθηση με την χρήση του sFlow σε OpenFlow switches κάναμε δοκιμές σε hardware (NEC IP8800) και σε software (Open vSwitch) switches που υποστήριζαν και τα δυο πρωτόκολλα (sFlow, OpenFlow). Το πρωτόκολλο μπορούσε να εξάγει τα δεδομένα παθητικής παρακολούθησης τα οποία αφορούσαν την κίνηση που δημιουργούσαν οι πειραματικοί κόμβοι των διαφόρων slices και στις δυο περιπτώσεις. Η επιβάρυνση της επίδοσης των OpenFlow switches και στις δυο περιπτώσεις δεν ήταν σημαντικές. Η ακρίβεια των μετρήσεων στην περίπτωση του NEC IP8800 εκτιμήθηκε βάσει των μετρητών *Dropped sFlow samples* και *Overflow Time of sFlow Queue* που διαθέτει η συγκεκριμένη υλοποίηση. Η αντίστοιχη εκτίμηση στο Open vSwitch έγινε συγκρίνοντας τον αριθμό των πακέτων που συλλέχθηκαν κατά την διάρκεια του πειράματος, με τον θεωρητικό αριθμό που περιμέναμε να συλλέξουμε και ήταν ίσος με το (sampling rate)x(total packet number).

Πραγματοποιήσαμε μετρήσεις χρησιμοποιώντας το εργαλείο Iperf και το Tcpreplay παρουσιάζοντας ενδεικτικά αποτελέσματα για το NEC IP8800. Τα δείγματα κίνησης που χρησιμοποιήθηκαν για τις δοκιμές με το Tcpreplay είχαν συλλεχθεί από το Τομέα Επικοινωνιών, Ηλεκτρονικής και Συστημάτων Πληροφορικής. Το sampling rate που ορίστηκε για τις δοκιμές ήταν (1/256). Οι επιδόσεις που αναγράφονται (**Πίνακας 5-1**) αντιστοιχούν με την μέγιστη κίνηση που μπορούσε να χειριστεί το OpenFlow switch διατηρώντας το μετρητή *Dropped sFlow samples* μηδενικό.

Κατά την εγκαθίδρυση νέων flow entries στο IP NEC8800 παρατηρήθηκε ότι όταν ο αριθμός των εγκαθιδρυμένων flows έφτανε το μέγιστο αριθμό που μπορούσε να υποστηρίξει το switch η λειτουργία του sFlow επηρεαζόταν αρνητικά. Δεν γινόταν συλλογή των πακέτων βάσει του sampling rate καθώς χάνονταν πακέτα (dropped packets). Το συγκεκριμένο hardware switch καθώς και αρκετά ακόμα σε αυτή την κατηγορία κρατούν ένα μέρος των flow entries σε μνήμες ειδικού τύπου TCAM [Gupt01] για μεγαλύτερη ταχύτητα επεξεργασίας των πακέτων και ένα άλλο μέρος

στην μνήμη γενικού σκοπού της συσκευής για οικονομία ενέργειας και μείωση του κόστους. Η ύπαρξη εγγραφών στη μνήμη γενικού σκοπού και οι ενέργειες που εκτελούνται πάνω σε αυτές (εισαγωγή, διαγραφή, ανάγνωση) αφενός μειώνουν τον ρυθμό μεταγωγής πακέτων και αφετέρου δημιουργούν προβλήματα στην συνολική λειτουργία διεργασιών της συσκευής που σχετίζονται και με άλλες λειτουργίες του επιπέδου ελέγχου και διαχείρισης της συσκευής.

	Iperf			tcpreplay	
PPS (packet/s)	290	214	114	290	360
CPU load (%)	20	32	19	26	28
MTU (bytes)	256	512	1024	-	-
B/W (Mbit/s)	560	848	919	-	-
Overflow time (s)	3	0	0	1	3

Πίνακας 5-1: Πίνακας μετρήσεων για τη λειτουργία του sFlow σε OpenFlow hardware switch NEC IP8800

Η αποφυγή χρήσης της κεντρικής μνήμης για αποθήκευση flow entries σε ένα διαμοιραζόμενο εικονικό περιβάλλον που βασίζεται πάνω σε μια τοπολογία από OpenFlow switches επιβάλλει τον περιορισμό ανά χρήστη του αριθμού των flow entries που μπορεί να εισάγει σε κάθε OpenFlow switch. Στη περίπτωση του PaFloMon ο συγκεκριμένος περιορισμός επιβάλλεται με ρύθμιση της παραμέτρου *number of flow entries per slice per dpid* στον FlowVisor, ο οποίος ως proxy controller ανάμεσα στου OpenFlow controllers των χρηστών και τα OpenFlow switches της υποδομής έχει την ικανότητα να επιβάλλει μια πολιτική όσον αφορά τον αριθμό των OpenFlow entries ανά switch ανά χρήστη.

5.6.2 Πειραματικές Μετρήσεις Μηχανισμού Δικτυακών Ανωμαλιών

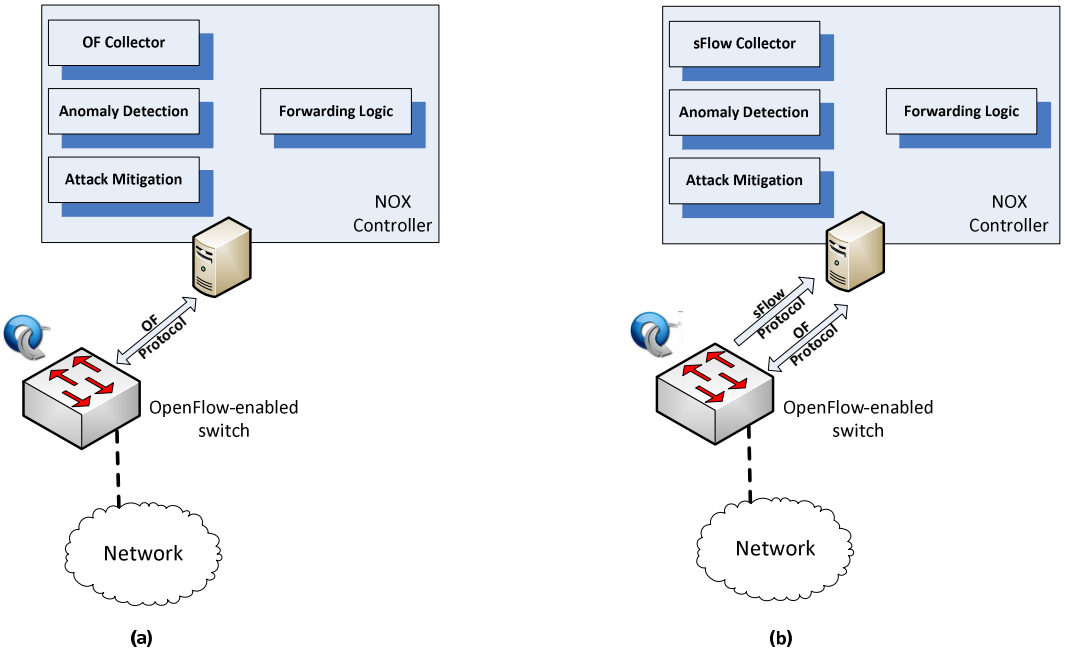
Ακολουθούν πειραματικές μετρήσεις που επιτρέπουν την σύγκριση μεταξύ ενός μηχανισμού παθητικής παρακολούθησης που χρησιμοποιεί μόνο το OpenFlow για την συλλογή και εξομάλυνση των δικτυακών ανωμαλιών σε σύγκριση με ένα μηχανισμό που κάνει χρήση του sFlow για την συλλογή των δεδομένων. Ο αλγόριθμος επεξεργασίας των δεδομένων βασίζεται σε μια μέθοδο ελέγχου της εντροπίας, η οποία

είναι ανεξάρτητη της τοπολογίας και μπορεί να χρησιμοποιηθεί για εντοπισμό και ταξινόμηση διαφορετικών ανωμαλιών [Shah07] [Andr09].

Στην περίπτωση της χρησιμοποιούμενης μεθόδου η μέτρηση της εντροπίας αποτυπώνει την τυχαιότητα των τιμών που έχουν συγκεκριμένα πεδία των επικεφαλίδων των πακέτων που λαμβάνονται ως δείγμα. Υψηλή εντροπία υποδηλώνει μια μεγαλύτερη διασπορά στη συνάρτηση πυκνότητας πιθανότητας, ενώ χαμηλή εντροπία υποδηλώνει την ύπαρξη επαναλαμβανόμενων τιμών στο δείγμα. Οι επικεφαλίδες που λαμβάνονται υπόψη για τον υπολογισμό των εντροπιών είναι η source IP address (srcIP), η destination IP address (dstIP), η source port (srcPort), η destination port (dstPort). Μετρήσεις πραγματοποιήθηκαν για κατανεμημένες επιθέσεις άρνησης υπηρεσίας (DDoS), αυτοδιαδιδόμενους ιούς (worm propagation) και σάρωση πορτών υπηρεσίας (port scan). Η ανάλυση των επιπτώσεων που έχει στην εντροπία η εκάστοτε επίθεση αναλύονται στο [Giot14].

Για την εισαγωγή κανόνων στα OpenFlow switches και την εξομάλυνση των ανωμαλιών που εντοπίζονταν με την μέθοδο ελέγχου της εντροπίας χρησιμοποιήσαμε τον NOX OpenFlow controller [NOX]. Ο NOX είναι ένα προγραμματιζόμενος controller που δίνει την δυνατότητα μέσω του API που διαθέτει να αλληλεπιδρά με άλλα προγράμματα (NOX applications). Τα 3 δομικά στοιχεία του μηχανισμού **Σχήμα 10** υλοποιήθηκαν ως NOX applications με την βοήθεια της γλώσσα προγραμματισμού Python. Τα δείγματα κίνησης που χρησιμοποιήθηκαν στο επίπεδο προώθησης για την διεξαγωγή πειραματικών μετρήσεων είχαν συλλεχθεί από το δίκτυο υπολογιστών του Εθνικού Μετσοβίου Πολυτεχνείου και ήταν της τάξεως των 50 Mbps και 100 Mbps και 500Mbps. Τα δείγματα κίνησης ήταν απαραίτητο να τα χειριστούν OpenFlow switches με υποστήριξη στο sFlow. Για αυτό το σκοπό χρησιμοποιήσαμε το Open vSwitch εγκατεστημένο σε ένα Dual-core 3 GHz με 8 GB μνήμη. Ο έλεγχος ικανότητας χειρισμού δικτυακής κίνησης που έγινε με την βοήθεια του NTOP [Deri00] έδειξε ότι το OVS μπορούσε να χειρίζεται αξιόπιστα κίνηση της τάξεως των 100 Mbps και ταυτόχρονα να δειγματοληπτεί με συχνότητα 1/64 ή να χειρίζεται 500 Mbps με συχνότητα δειγματοληψίας πακέτων 1/256. Για κίνηση που έφτανε τα 500 Mbps πραγματοποιήσαμε τα πειράματά μας και στον hardware μεταγωγέα NEC IP8800/S3640.

Η πειραματική διάταξη περιγράφεται στο **Σχήμα 11,** με την διάταξη εντοπισμού και εξομάλυνσης ανωμαλιών να χρησιμοποιεί μόνο το OpenFlow στην (a) περίπτωση και να συνδυάζει το OpenFlow και το sFlow στην (b) περίπτωση. Για τις ανάγκες των πειραματικών μετρήσεων και την αξιολόγηση των δυνατοτήτων των δυο λύσεων δημιουργήσαμε κίνηση που την κατευθύναμε τόσο στο OVS όσο και στο NEC switch. Για χαμηλού ρυθμούς ροής δεδομένων συγκρίναμε τις δυο λύσεις χρησιμοποιώντας στην περίπτωση (a) τον OF collector και στην περίπτωση (b) τον sFlow collector. Στην περίπτωση του OF collector, απαιτήθηκε η παραμετροποίηση της τιμής του soft-timeout των flow entries στα OpenFlow switches, ώστε να προλαβαίνει ο OF collector να συλλέγει τα στατιστικά στοιχεία για τα flow entries πριν κάνει την διαγραφή τους το OF switch.



Σχήμα 11: Διάταξη εντοπισμού και εξομάλυνσης ανωμαλιών με (a) χρήση OpenFlow (b) συνδυαστική χρήση OpenFlow και sFlow

Δεδομένου ότι το παράθυρο συλλογής δεδομένων ήταν τα 30 δευτερόλεπτα, ορίσαμε το soft-timeout στα 31 δευτερόλεπτα. Στην περίπτωση του sFlow collector η υλοποίηση έγινε μέσα στον NOX controller. Το συγκεκριμένο application του NOX συμπληρωνόταν με ακόμα ένα NOX application που ήταν υπεύθυνο για την παραγωγή

κανόνων προώθησης δεδομένων ακολουθώντας την λογική του MAC learning and forwarding, που υλοποιούν τα κλασσικά switches.

Τα δείγματα κίνησης που είχαμε συλλέξει από το Ε.Μ.Π χρησιμοποιήθηκαν για τη δημιουργία κίνησης, την οποία καλούνταν να χειριστούν τα OpenFlow switches της πειραματικής διάταξης. Το εργαλείο που έκανε αναδημιουργία της κίνησης από το συλλεγμένο δείγμα ήταν το Tcpreplay [Tcpreplay]. Η δημιουργία των πακέτων, που προσομοίωναν τον εκάστοτε τύπο επίθεσης, παρήχθησαν με την βοήθεια του εργαλείου Scapy [Scapy].

Για την εξομοίωση μιας DDoS επίθεσης δημιουργήσαμε πακέτα TCP/SYN με συγκεκριμένες dstIP και dstPort και τυχαία επιλεγμένες srcIP και srcPort τιμές. Για την επίθεση τύπου worm propagation χρησιμοποιήσαμε την προαναφερθείσα τακτική με την διαφορά ότι η dstIP και dstPort ήταν τυχαία επιλεγμένες, ενώ η srcIP και srcPort ήταν συγκεκριμένες. Τέλος, για την δημιουργία δείγματος πακέτων επίθεσης τύπου port scanning, η dstIP και η srcIP ήταν καθορισμένες, ενώ οι dstPort και οι srcPort ήταν τυχαίες.

Οι τιμές των αριθμητικών παραμέτρων στα σενάρια των πειραματικών μετρήσεων που εκτελέσαμε συνοψίζονται (**Πίνακας 5-2).** Όσο μεγαλύτερος είναι ο ρυθμός κίνησης τόσο μικρότερο ρυθμό δειγματοληψίας επιλέξαμε να έχουμε, όπως φαίνεται από τον πίνακα. Ο ρυθμός πακέτων της ροής επίθεσης αυξάνει όσο αυξάνει ο μέσος ρυθμός κίνησης.

Ο ρυθμός δειγματοληψίας μειώθηκε σε μεγάλους ρυθμούς κίνησης καθώς δεν είναι δυνατή η εξαγωγή στατιστικών με τη χρήση sFlow, αφού κάτι τέτοιο απαιτεί μεγάλη επεξεργαστική ισχύ στα switches, η οποία δεν είναι διαθέσιμη. Με στόχο τον εντοπισμό ανωμαλιών που εντοπίζονται και χαρακτηρίζουν μεγάλους όγκους κίνησης (π.χ DDoS), αυξάνουμε τον ρυθμό επίθεσης, θεωρώντας το OpeFlow switch σημείο συνάθροισης κίνησης.

	Μέσος ρυθμός κίνησης (Mbps)	**Δειγματοληψία**		**Ρυθμός Επίθεσης (packets/sec)**
Σενάριο 1	50	No	1/64	50 - 200
Σενάριο 2	100	1/64		200-500
Σενάριο 3	500	1/256		1000-2500

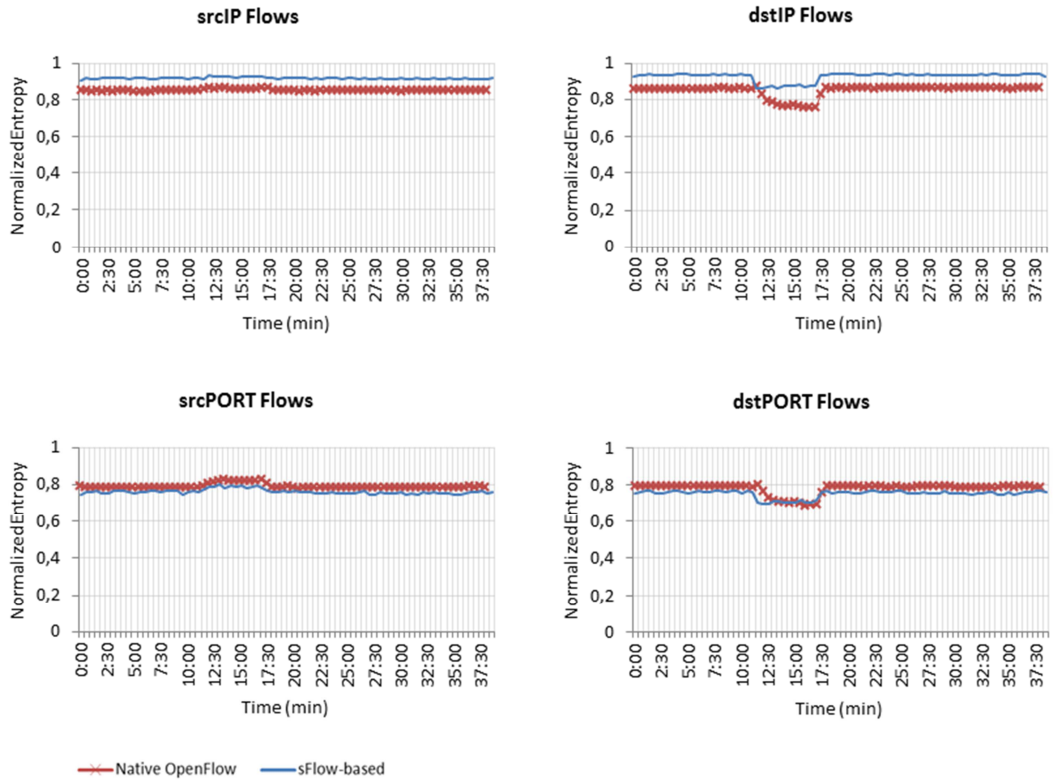
Πίνακας 5-2: Τιμές αριθμητικών παραμέτρων των πειραματικών μετρήσεων της διάταξης εντοπισμού και εξομάλυνσης ανωμαλιών

Τα πειραματικά σενάρια μπορεί να συνοψιστούν ως εξής:

- 50 Mbps κίνησης κατευθυνόμενα στην διάταξη (a) και (b), **Σχήμα 11,** για την εξαγωγή συγκρίσιμών αποτελεσμάτων.

Ο ρυθμός δειγματοληψίας του sFlow ήταν 1/64. Χρησιμοποιήθηκε και στις δυο (2) περιπτώσεις OVS ως OpenFlow switch. Σκοπός μας ήταν να διαπιστώσουμε ότι η sFlow μέθοδος που χρησιμοποιεί δειγματοληψία μπορεί αξιόπιστα να εντοπίσει ανωμαλίες. Ο ρυθμός των πακέτων που αντιπροσώπευαν την επίθεση ήταν 50-200 πακέτα/δευτερόλεπτο, αριθμός μικρός συγκρινόμενο με τα 12.000-13.000 πακέτα/δευτερόλεπτο της κανονικής κίνησης.

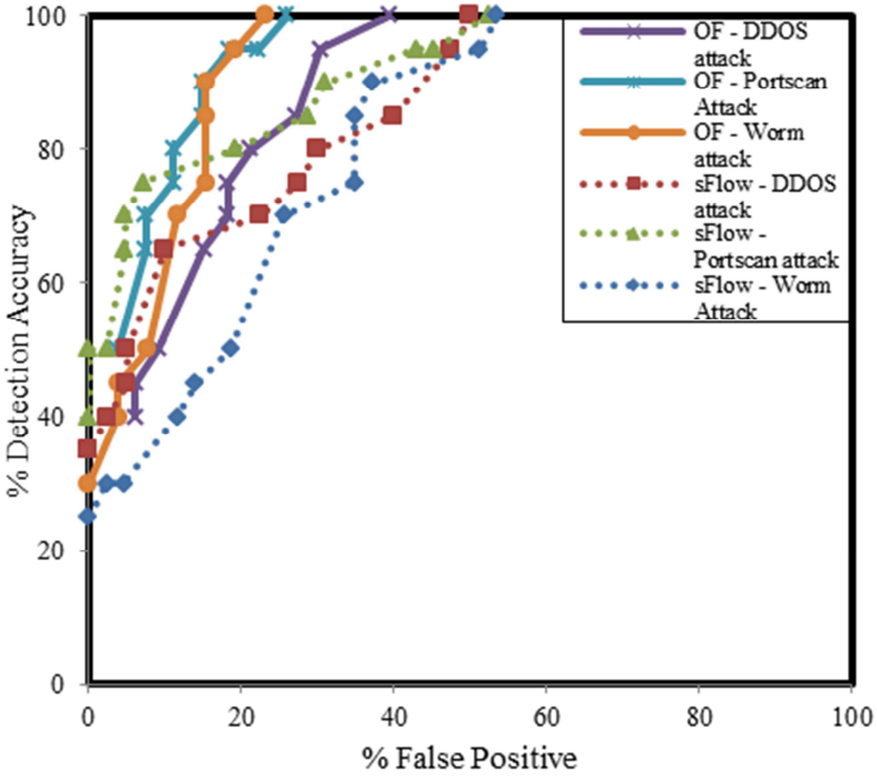
- 100 Mbps κίνησης κατευθυνόμενα στην διάταξη (b), **Σχήμα 11**. Ο ρυθμός δειγματοληψίας του sFlow ήταν 1/64. Χρησιμοποιήθηκε το OVS ως OpenFlow switch.
- 500 Mbps κίνησης κατευθυνόμενα στην διάταξη (b) του Σχήματος 11. Ο ρυθμός δειγματοληψίας του sFlow ήταν 1/256. Χρησιμοποιήθηκε το OVS και το NEC IP8800 ως OpenFlow switch.



Σχήμα 12: Τιμές εντροπίας υπολογισμένες με χρήση OpenFlow και sFlow στατιστικών στοιχείων

Η σύγκριση των αποτελεσμάτων των δυο (2) μεθόδων για 50 Mbps στον υπολογισμό της εντροπίας τεσσάρων χαρακτηριστικών μεγεθών (srcIP, dstIP, srcPort, dstPort) φαίνεται στο **Σχήμα 12** κατά τη διάρκεια μιας DDoS επίθεσης έξι λεπτών. Μπορούμε να παρατηρήσουμε ότι τα αποτελέσματα της sFlow μεθόδου είναι ικανοποιητικά, αφού αποτυπώνουν την μείωση της εντροπίας σε dstIP και dstPort, με αποτέλεσμα να είναι δυνατός ο εντοπισμός της επιθέσεως και μέσω της sFlow μεθόδου. Μπορούμε επίσης να παρατηρήσουμε μια ελαφριά άνοδο στην εντροπία της srcPort, που αποτυπώνεται και στις 2 μεθόδους.

Παρακάτω παρουσιάζουμε τις καμπύλες Receiver Operating Characteristic (ROC) για τα ποσοστά εντοπισμού αληθών συναγερμών (true positive) και ψευδών συναγερμών (false positive) όσον αφορά τα μίγματα ροών (**Πίνακας 5-2)**.



Σχήμα 13: ROC καμπύλες ανίχνευσης για DDoS, Worm και Portscan επιθέσεις με αποκλειστική χρήση OpenFlow και συνδυαστική χρήση OpenFlow και sFlow για 50 Mbps

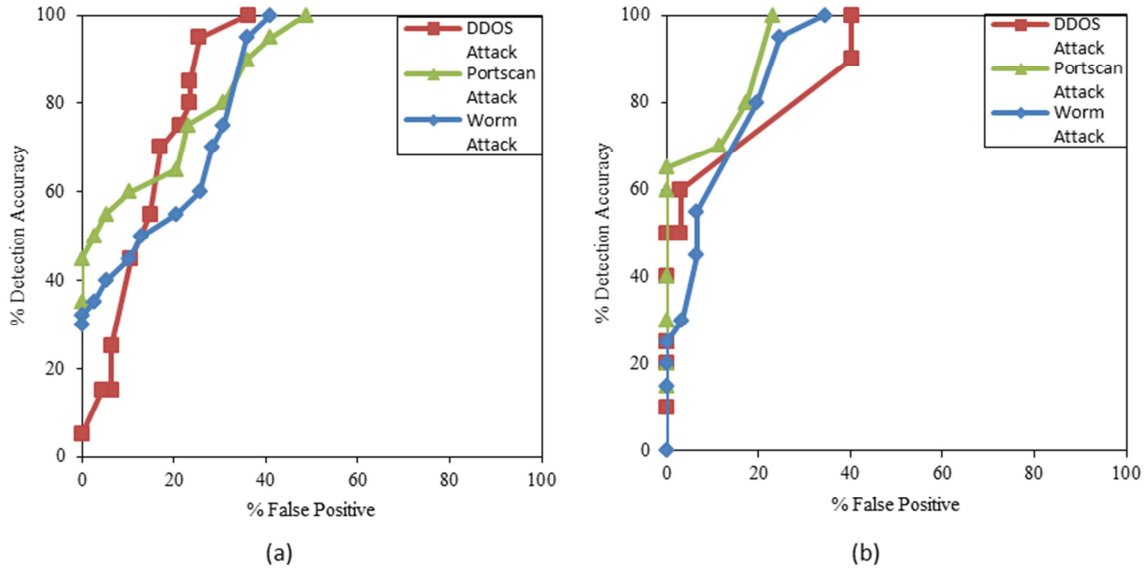
Το **Σχήμα 13** δείχνει ROC καμπύλες για DDoS, Worm και Portscan επιθέσεις με αποκλειστική χρήση OpenFlow και συνδυαστική χρήση OpenFlow και sFlow. Όπως παρατηρούμε, η υλοποίηση με αποκλειστική χρήση OpenFlow έχει καλύτερη απόδοση καθώς δεν υπάρχει δειγματοληψία και το σύνολο των ροών έχει καταγραφεί και μπορεί να τύχει επεξεργασίας από το αντίστοιχο δομικό στοιχείο που έχουμε παρουσιάσει παραπάνω. Η μέθοδος του sFlow έχει ελαφρά χειρότερη επίδοση, με την απαίτηση για 100% επιτυχία στον εντοπισμό να δημιουργεί περίπου 50% λανθασμένους συναγερμούς. Η sFlow έχει το πλεονέκτημα του μικρού φόρτου που δημιουργεί στο CPU του switch, με τα συγκριτικά αποτελέσματα να παρουσιάζονται παρακάτω.

Η κλιμακοθετησιμότητα της OpenFlow υλοποίησης δεν δημιούργησε προβλήματα με ταχύτητες μέχρι τα 100 Mbps. Σε υψηλότερους ρυθμούς (π.χ 100-500 Mbps) η περιοδική συλλογή και επεξεργασία των στατιστικών με χρήση του OpenFlow δεν λειτουργούσε με ομαλό τρόπο. Ο χρόνος επεξεργασίας που καταναλωνόταν για την εξαγωγή των στατιστικών ήταν μεγάλος και συγχρόνως παρατηρήθηκαν

καθυστερήσεις στην προώθηση των πακέτων στο επίπεδο προώθησης δεδομένων. Ορισμένες μάλιστα DDoS επιθέσεις δημιούργησαν πολλές ταυτόχρονες νέες ροές, με αποτέλεσμα, τα ερωτήματα χειρισμού που έκανε το OpenFlow switch στον OpenFlow controller για τις συγκεκριμένες ροές να δημιουργούν υπερφόρτωση του επιπέδου ελέγχου που αποτυπώθηκε στο CPU load των δυο συσκευών.

Στο **Σχήμα 14** παρουσιάζουμε διάγραμμα με ROC καμπύλες ανίχνευσης για DDoS, Worm και Portscan επιθέσεις για την υλοποίηση που κάνει συνδυαστική χρήση OpenFlow και sFlow για (a) 100 Mbps και (b) 500 Mbps. Στην περίπτωση των 100 Mbps η κανονική κίνηση ήταν της τάξεως των 25.000 πακέτων/δευτερόλεπτο και η επίθεση 200-500 πακέτων/δευτερόλεπτο. Όπως φαίνεται στην (a) περίπτωση η μέθοδος του sFlow για 100% επιτυχία στον εντοπισμό δημιουργεί περίπου 40%, 42% και 50% λανθασμένους συναγερμούς για DDoS, Worm propagation και port scanning επιθέσεις αντίστοιχα. Στην (b) περίπτωση, των 500 Mbps, η κανονική κίνηση ήταν της τάξεως των 130.000 πακέτων/δευτερόλεπτο και η επίθεση 1000-2500 πακέτων/δευτερόλεπτο. Όπως φαίνεται στην (b) περίπτωση η μέθοδος του sFlow για 100% επιτυχία στον εντοπισμό δημιουργεί περίπου 23%, 27% και 34% λανθασμένους συναγερμούς για DDoS, Worm propagation και port scanning επιθέσεις αντίστοιχα. Συμπεραίνουμε λοιπόν ότι για μεγαλύτερους ρυθμούς ροής δεδομένων και τηρουμένης της αναλογίας μεταξύ πακέτων κανονικής κίνησης και κίνησης επιθέσεως η μέθοδος sFlow, με δειγματοληψία 1/256 πετυχαίνει βελτιωμένα αποτελέσματα.

Οι μεγάλοι ρυθμοί κίνησης, σε περιβάλλοντα δικτύων με χιλιάδες hosts, ισοδυναμούν με μεγαλύτερο εύρος τιμών στις επικεφαλίδες των πακέτων και άρα μεγαλύτερή εντροπία. Το επακόλουθο είναι ότι η μέθοδος εντοπισμού με χρήση της εντροπίας αποδίδει καλύτερα σε περιβάλλοντα που η κίνηση προέρχεται από πολλές διαφορετικές πηγές και κατευθύνεται σε πολλούς διαφορετικούς παραλήπτες, όπως είναι τα δίκτυα των παρόχων υπηρεσιών Internet. Οι τιμές που επιτύχαμε με την χρήση του NEC IP8800 ήταν συγκρίσιμες με αυτές του OVS, γεγονός που δείχνει ότι το software switch δεν αντιμετώπισε πρόβλημα με τον χειρισμό των πακέτων ή με το sFlow sampling σε ταχύτητες της τάξεως των 500 Mbps.



Σχήμα 14: ROC καμπύλες ανίχνευσης για DDoS, Worm και Portscan επιθέσεις με συνδυαστική χρήση OpenFlow και sFlow για (a) 100 Mbps και (b) 500 Mbps

Η παθητική παρακολούθηση των ροών από τα OpenFlow switches καταναλώνει πόρους, με τους βασικότερους να είναι η κατανάλωση επεξεργαστικής ισχύος σε Openflow switches και controller και ο αριθμός των flow entries (το τελευταίο μόνο στην περίπτωση χρήσης OpenFlow για παρακολούθηση). Μετρούμε το όφελος που υπάρχει στην χρήση των συγκεκριμένων πόρων με χρήση της συνδυαστικής μεθόδου (OpenFlow και sFlow) για κανονική κίνηση και κίνηση που περιλαμβάνει ροή επίθεσης 200 πακέτων/δευτερόλεπτο. Όπως παρατηρούμε (**Πίνακας 5-3),** κατά την διάρκεια της επίθεση το sFlow συντελεί στην μείωση της χρήσης του CPU από 61% σε 39% και σε αντίστοιχη μείωση του αριθμού των μέσων ροών από 7685 σε 351. Η διαφορά είναι μεγάλη, ακόμα και στην περίπτωση που δεν υπάρχει ροή επίθεσης μέσα στο δείγμα κίνησης που χειρίζεται το OpenFlow switch, με το 47% χρήσης CPU και το μέσο αριθμό των 5184 εγγραφών στο OpenFlow table να μειώνεται σε 32% και 217 αντίστοιχα με την βοήθεια του sFlow.

Αξίζει να σημειώσουμε ότι οι 217 εγγραφές του OpenFlow table αντιστοιχούν σε κανόνες προώθησης πακέτων που έχουν επιβληθεί για το συγκεκριμένο σκοπό, ενώ οι 5184, για το ίδιο ακριβώς δείγμα κίνησης, αφορούν εγγραφές που ουσιαστικά επιτελούν το έργο της προώθησης πακέτων αλλά και το έργο της παθητικής παρακολούθησης. Το ίδιο ισχύει και στην διαφορά που παρατηρείται για το δεύτερο

δείγμα κίνησης που χρησιμοποιήθηκε και δημιουργεί την αύξηση των 351 εγγραφών προώθησης σε 7685 που απαιτούνται για την ταυτόχρονη προώθηση και παρακολούθηση.

Μέθοδος ανίχνευσης με μέτρηση εντροπίας	**50Mbps κίνηση χωρίς επίθεση**		**50Mbps κίνηση με 200 πακέτα/sec επίθεση**	
	Μέσος αριθμός ροών	**Χρήση CPU**	**Μέσος αριθμός ροών**	**Χρήση CPU**
(a) OpenFlow	5184	47%	7685	61%
(b) OpenFlow + sFlow	217	32%	351	39%

Πίνακας 5-3: Επιβάρυνση στην χρήση CPU του OpenFlow switch και στο μέσο αριθμό κανόνων με/χωρίς επίθεση για παθητική παρακολούθηση με (a) OpenFlow και (b) sFlow

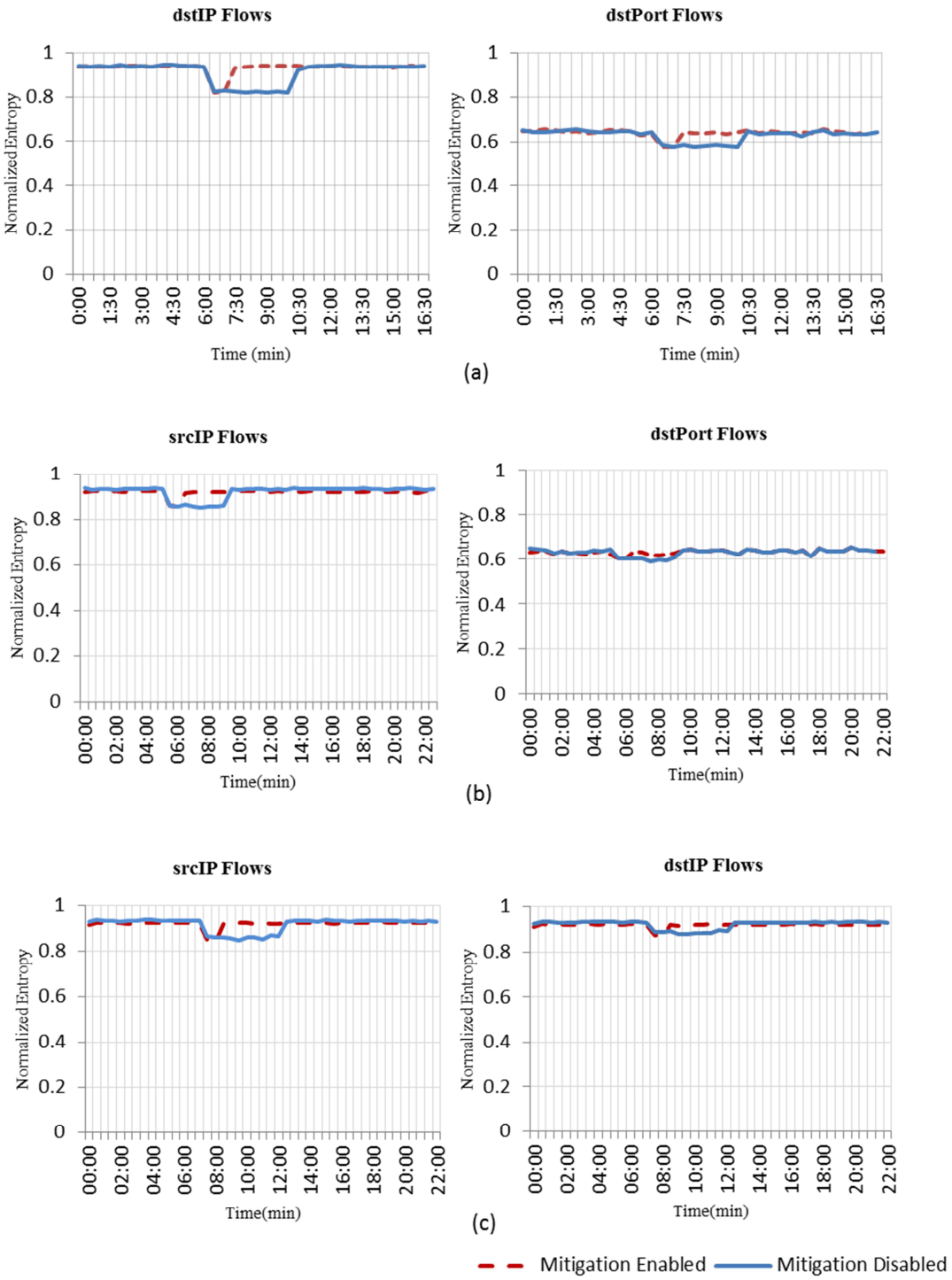
Εντατική χρήση πόρων συντελείται και στην πλευρά του OpenFlow controller, όταν γίνεται αποκλειστική χρήση OpenFlow, καθώς η μορφή της κίνησης στο επίπεδο προώθησης δεδομένων επηρεάζει τον αριθμό των ερωτημάτων που πραγματοποιούνται στον controller. Όταν το switch δεν έχει εγκαθιδρυμένες ροές που να αντιστοιχούν με τα πακέτα που επεξεργάζεται δημιουργεί ερωτήματα προς τον OpenFlow controller. Στην πλευρά του OpenFlow controller βλέπουμε ότι για το χειρισμό του δείγματος κίνησης χωρίς επίθεση, το sFlow μειώνει την χρήση του CPU από 42% σε 25% και κατά τη διάρκεια επίθεσης από 63% σε 31%.

Μέθοδος ανίχνευσης με μέτρηση εντροπίας	**50Mbps κίνηση χωρίς επίθεση**	**50Mbps κίνηση με 200 πακέτα/sec επίθεση**
OpenFlow (a)	42%	63%
OpenFlow + sFlow (b)	25%	31%

Πίνακας 5-4: Επιβάρυνση στην χρήση CPU του OpenFlow controller με/χωρίς επίθεση για παθητική παρακολούθηση με (a) OpenFlow και (b) sFlow

Μπορούμε συνεπώς να εξάγουμε το συμπέρασμα ότι η συνδρομή του sFlow είναι σημαντική στην μείωση χρήσης πολύτιμων πόρων στα OpenFlow switches και συγχρόνως στους OpenFlow controllers.

Το τελευταίο βήμα του μηχανισμού που έχουμε περιγράψει για τον εντοπισμό των ανωμαλιών στη δικτυακή κίνηση περιλαμβάνει την εξομάλυνσή τους. Χρησιμοποιώντας το δείγμα κίνησης, όπως στο **Σχήμα 12,** παρουσιάζουμε τις τιμές της εντροπίας κατά την διάρκεια (a) DDoS, (b) worm propagation, (c) port scanning με και χωρίς τον μηχανισμό εξομάλυνσης σε λειτουργία. Παρατηρούμε ότι μετά την εγκαθίδρυση των κατάλληλων flow entries που απορρίπτουν τα πακέτα που αντιστοιχούν σε ροές επίθεσης η εντροπία της συνολικής κίνησης επιστρέφει στις αναμενόμενες τιμές που υπήρχαν πριν την έναρξή τους. Χαρακτηριστικότερο παράδειγμα αποτελεί αυτό της DDoS, δεδομένου ότι εκεί η εντροπία έχει μεταβληθεί, λόγω επίθεσης. Στην συγκεκριμένη περίπτωση η αποτελεσματική αντιμετώπιση της επίθεσης επιβάλλει την απόρριψη ροών που κατευθύνονται στο θύμα της επίθεσης, καθώς οι πηγές επίθεσης μπορεί να είναι χιλιάδες και άρα είναι κοστοβόρο, από άποψη πόρων (εγγραφές ροών, CPU), η εγκαθίδρυση κανόνων για την απόρριψη χιλιάδων ροών με κριτήριο τις διαφορετικές πηγές από τις οποίες έχει ξεκινήσει η επίθεση ταυτόχρονα.



Σχήμα 15: Τιμές της εντροπίας κατά την διάρκεια (a) DDoS, (b) worm propagation, (c) port scanning με/χωρίς τον μηχανισμό εξομάλυνσης

6 Ανάθεση Ροών ανά Εικονικό Δίκτυο σε Υποδομές Οριζόμενες από Λογισμικό στο Επίπεδο Ελέγχου

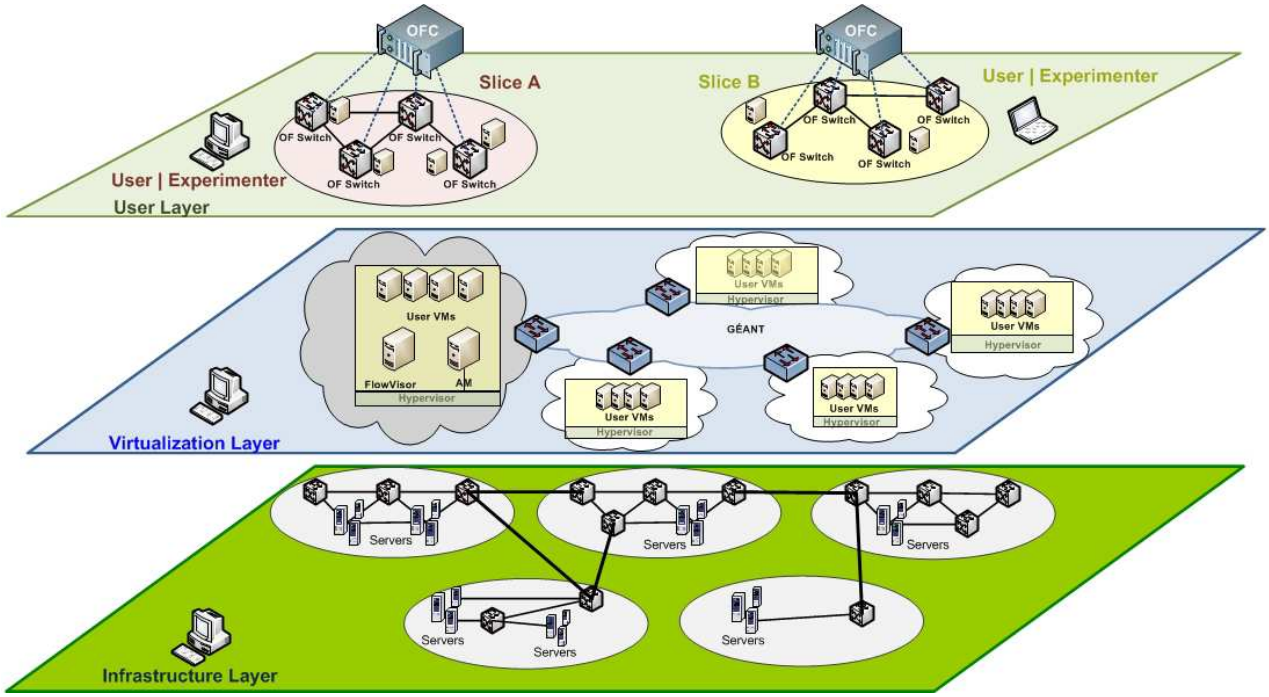
Network Virtualization in Multi-Tenant Software Defined Networking

6.1 Ερευνητικός Στόχος

Ο διαχωρισμός του επιπέδου ελέγχου και του επιπέδου προώθησης δεδομένων με χρήση του OpenFlow καθώς και οι τεχνολογίες εικονικοποίησης δημιουργούν το κατάλληλο υπόβαθρο για νέου τύπου υπηρεσίες. Είναι δυνατή η διασύνδεση πειραματικών υποδομών και η απόδοση του δικτυακού ελέγχου στους χρήστες, η ανάπτυξη νέων υπηρεσιών διασυνδεδεμένων IaaS (Networked IaaS) σε περιβάλλοντα cloud computing αλλά και ανάπτυξη νέων υπηρεσιών σε δίκτυα ευρείας περιοχής (Wide Area Network – WAN).

Εταιρείες με υπηρεσίες παγκοσμίου κλίμακας άρχισαν να αναπτύσσουν ιδιοταγή δίκτυα οριζόμενα από λογισμικό (Software Defined Networks – SDNs) πολύ μεγάλης κλίμακας (global-scale) με υψηλή ποιότητα υπηρεσίας (production-quality), εκμεταλλευόμενες το πρωτόκολλο OpenFlow. Χαρακτηριστικά παραδείγματα τέτοιου τύπου αποτελούν το εσωτερικό WAN δίκτυο που ανέπτυξε η Google μεταξύ των datacenters που διαθέτει σε όλες τις ηπείρους [Kopo10] καθώς και το IaaS της NTT communications [NTT], του μεγαλύτερου τηλεπικοινωνιακού παρόχου παγκοσμίως. Στις συγκεκριμένες υλοποιήσεις, αν και το στοιχείο της καινοτομίας είναι έντονο σε επίπεδο αρχιτεκτονικής σχεδίασης και υλοποίησης, ο έλεγχος παραμένει στον πάροχο της υποδομής, με τον χρήστη να εκμεταλλεύεται προκαθορισμένες μόνο υπηρεσίες δικτύου.

Η τάση για δημιουργία μεγάλων υποδομών (infrastructure substrate) για απαιτητικούς χρήστες, με ανάγκες ελέγχου υπολογιστικών και δικτυακών πόρων για τη δημιουργία υπηρεσιών προστιθέμενης αξίας, δεν καλύπτεται από τις παραπάνω καινοτόμες προσπάθειες.



Σχήμα 16: Παρουσίαση του επιπέδου των φυσικών πόρων του δικτύου, του επιπέδου εικονικοποίησης/αφαίρεσης (ενδιάμεσο) και του χρήστη (υψηλότερο)

Το ζήτημα της παροχής λειτουργιών του επιπέδου ελέγχου σε πολλαπλούς χρήστες εικονικών δικτύων **Σχήμα 16)** παραμένει ανοιχτό.

Σε τέτοιες υποδομές θα πρέπει να υπάρχει η δυνατότητα ταυτόχρονης απόδοσης του ελέγχου εύρους δικτυακών ροών (flowspace) σε διαφορετικούς χρήστες, με διατήρηση της απομόνωσης. Ο χρήστης θα πρέπει να έχει την ικανότητα να επιβάλλει την δική του λογική δικτυακής διασύνδεσης, οριζόμενη από λογισμικό (SDN) μόνο στους πόρους που του έχουν αποδοθεί.

Η ανάθεση μέρους του flowspace σε χρήστες απαιτεί την ικανότητα ελέγχου πιθανών συγκρούσεων μεταξύ αιτημάτων διαφορετικών χρηστών (flowspace conflict detection) καθώς και την ανάλυση των αιτημάτων ως προς την συμβατότητα τους με την λογική διαχωρισμού (slicing) που θέλει να επιβάλλει ο διαχειριστής της υποδομής (flow analysis) [FOAM].

Χαρακτηριστικό παράδειγμα διαμοιρασμού δικτυακών πόρων σε ένα SDN δίκτυο, είναι ομαδοποίηση των ροών βάσει του VLAN ID των πακέτων και η απόδοση του ελέγχου κάθε τέτοιας ομάδας σε OpenFlow controller χρήστη. Με αυτό τον τρόπο ο

κάθε χρήστης έχει το δικό του OpenFlow controller και μπορεί να ελέγχει στο σύνολο της υποδομής όλες τις ροές πακέτων που έχουν το συγκεκριμένο VLAN ID (domain-wide VLAN-based slicing). Αντίστοιχα, ο λογικός διαχωρισμός των ροών που ελέγχονται ανά χρήστη μπορεί να βασίζεται σε άλλα κριτήρια, όπως για παράδειγμα, το εύρος των MAC addresses (MAC-based slicing), χρήσιμο σε περιβάλλοντα cloud computing.

Η ερευνητική μας εργασία έχει ως σκοπό να προσδιορίσει και να αξιολογήσει πολιτικές διαμοιρασμού του flowspace που εξασφαλίζουν την απομόνωση των πόρων που διαχειρίζονται οι χρήστες και αποτελούν τα εικονικά τους δίκτυα.

Επιγραμματικά αναφέρουμε τους άξονες της ερευνητικής μας εργασίας, όσον αφορά τους μηχανισμούς ανάθεσης ροών ανά εικονικό δίκτυο χρήστη σε περιβάλλοντα SDN κεντρικοποιημένου ελέγχου που κάνουν χρήση του πρωτοκόλλου OpenFlow:

- Αναλύουμε δυο (2) προσεγγίσεις για την υλοποίηση αρχιτεκτονικών δικτύωσης οριζομένων από λογισμικό, που εξυπηρετούν πολλαπλούς χρήστες και επιτρέπουν την απόδοση λειτουργιών ελέγχου στους χρήστες/ιδιοκτήτες (users/tenants) των εικονικών δικτύων. Η πρώτη αναφέρεται ως **Αρχιτεκτονική Proxy Controller** και η δεύτερη αναφέρεται ως **Αρχιτεκτονική Network Hypervisor [**Casa10**].**

- Ορίζουμε τρεις (3) διαφορετικές μεθόδους διαχωρισμού του flowspace που βρίσκουν εφαρμογή και στις δυο προαναφερθείσες αρχιτεκτονικές προσεγγίσεις:

  (α) **domain-wide slicing**

  Με την πρώτη μέθοδο ο χρήστης ελέγχει τις ροές πακέτων που ανήκουν στο κομμάτι του flowspace που του έχει αποδοθεί σε ολόκληρη την τοπολογία υποδομής.

  (β) **switch-wide slicing**

Με τη δεύτερη μέθοδο ελέγχει τις ροές πακέτων που ανήκουν στο κομμάτι του flowspace που του έχει αποδοθεί μόνο σε ένα υποσύνολο των OpenFlow switches της συνολικής τοπολογίας .

(γ) **port-wide slicing**

Στην τρίτη περίπτωση ελέγχει τις ροές πακέτων που ανήκουν στο κομμάτι του flowspace που του έχει αποδοθεί μόνο σε ένα υποσύνολο των διαθέσιμων πορτών των OpenFlow switches.

- Περιγράφουμε πως επιτυγχάνεται η απομόνωση βάσει των τριών μεθόδων και προτείνουμε τρόπους μείωσης των κανόνων που απαιτούνται για την εφαρμογή της πολιτικής διαμοιρασμού των πόρων, όπου αυτό είναι δυνατό.

- Υλοποιούμε μια μηχανή διαμοιρασμού του flowspace (Flowspace slicing policy -FSP) που δύναται να εφαρμόζει μια από τις τρεις μεθόδους. Η συγκεκριμένη μηχανή είναι υπεύθυνη για να μετατρέπει τις πολιτικές του επιπέδου διαχείρισης της υποδομής (substrate management-plane policies) σε κανόνες ελέγχου ροών των εικονικών δικτύων και χρησιμοποιείται για αξιολόγηση των μεθόδων.

- Αξιολογούμε την επίδοση των προτεινόμενων μεθόδων διαμοιρασμού με χρήση της μηχανής FSP πάνω σε πραγματικές τοπολογίες δικτύων, όπως είναι το Internet2/OS3E [IN2topo] και το GÉANT [GEANTtopo], με παραμέτρους σύγκρισης τους πόρους που χρησιμοποιούνται στο επίπεδο ελέγχου. Συγκεκριμένα, με τον αριθμό των κανόνων που απαιτούνται για την εφαρμογή της πολιτικής και τον λόγο αποδοχής αιτημάτων χρηστών προς το συνολικό αριθμό αιτημάτων που δέχεται μια υποδομή.

6.2 Ανασκόπηση της Υπάρχουσας Κατάστασης

Η πρώτη προσπάθεια διαχωρισμού σε διακριτές περιοχές του επιπέδου ελέγχου και απόδοσης των επιμέρους περιοχών σε χρήστες δικτύων οριζομένων από λογισμικό έγινε με το σχεδιασμό του OpenFlow proxy controller. Ο OpenFlow proxy controller, με πρωτότυπο παράδειγμα αυτό του FlowVisor [Sher10], δημιουργεί ένα ενδιάμεσο επίπεδο ελέγχου.

Ο FlowVisor επέτρεπε τον έλεγχο του επιπέδου προώθησης από τους χρήστες (οι χρήστες εφήρμοζαν τη λογική προώθησης που αυτοί ήθελαν να επιβάλλουν), αλλά δεν επέτρεπε τη δημιουργία στρώματος αφαίρεσης δικτύου ανάμεσα στην πραγματική υποδομή και στο εικονικό δίκτυο του χρήστη.

Ο FlowVisor δεν είχε την ικανότητα δημιουργίας πολύπλοκων λειτουργιών, όπως είναι: (α) η υλοποίηση μιας εικονικής διαδρομής με δυνατότητα ευέλικτης αντιστοίχισης στις ζεύξεις υποδομής (path migration) ή (β) την υλοποίηση μια εικονικής διαδρομής πάνω από πολλαπλές οδεύσεις της τοπολογίας υποδομής (path splitting) [Yu08]. Επίσης ο κάθε χρήστης είχε πρόσβαση σε λειτουργίες, όπως είναι η ανακάλυψη τοπολογίας δικτύου (network discovery), που εξέθεταν την τοπολογία υποδομής. Παρά τους περιορισμούς στο σχεδιασμό του, έτυχε μεγάλης αποδοχής από την ερευνητική κοινότητα και συνεχίζει να εξελίσσεται και να χρησιμοποιείται από πολλές ερευνητικές υποδομές (π.χ. GENI, OFELIA, GÉANT OpenFlow testbed).

Στην ίδια λογική με τον FlowVisor άρχισαν να αναπτύσσονται και άλλοι OpenFlow proxy controllers, όπως είναι ο Flowspace Firewall [FlowFire], μια δημιουργία του Global Research Network Operations Center (GRNOC) [GRNOC]. Ο σχεδιασμός του συγκεκριμένου OpenFlow proxy controller εστιάζει στην βελτίωση της απόδοσης σε επίπεδο κλίμακας λειτουργίας και μείωση της χρονικής επιβάρυνσης που εισάγει ως ενδιάμεσο επίπεδο ελέγχου, με αντάλλαγμα για την βελτίωση των επιδόσεων την υποστήριξη τμηματοποίησης του επιπέδου ελέγχου μόνο στο Layer 2 της στοίβας πρωτοκόλλων. Οι προτεινόμενες μέθοδοι διαμοιρασμού του επιπέδου ελέγχου που μελετώνται στα πλαίσια της ερευνητικής μας εργασίας, με χρήση της μηχανής FSP, χρησιμοποιούν παραμέτρους Layer 2 και άρα είναι συμβατές και με τις δυο παραπάνω υλοποιήσεις OpenFlow proxy controller καθώς και με αρχιτεκτονικές που υποστηρίζουν network abstractions (path splitting, path migration).

Η αποδοτική αντιστοίχιση των εικονικών κόμβων και ζεύξεων που απαρτίζουν τα εικονικά δίκτυα και περιγράφονται στις αιτήσεις των χρηστών σε πραγματικούς πόρους της υποδομής είναι προαπαιτούμενο ενός δυναμικού περιβάλλοντος πολλαπλών χρηστών. Το συγκεκριμένο πρόβλημα που αναφέρεται ως πρόβλημα ενσωμάτωσης των εικονικών δικτύων (virtual network embedding problem – VNE) έχει μελετηθεί στην βιβλιογραφία [Chow12] [Papa13]. Ενδιαφέρον παρουσιάζει η διερεύνηση μεθόδου εφαρμογής στο δίκτυο υποδομής των αποτελεσμάτων του VNE αλγορίθμου για την αποδοτική αντιστοίχιση των εικονικών κόμβων και ζεύξεων που απαρτίζουν τα εικονικά δίκτυα. Η εφαρμογή της λύσης απαιτεί την δημιουργία κανόνων οι οποίοι εξασφαλίζουν την απομόνωση των χρηστών καθώς και την απόδοση των πόρων. Οι τρεις (3) διαφορετικές μέθοδοι διαχωρισμού του flowspace βρίσκουν εφαρμογή σε κάθε περίπτωση, αφού αποτελούν τον μηχανισμό υλοποίησης του αποτελέσματος που έχει προκύψει από την επίλυση του VNE προβλήματος.

Μια από τις πρώτες προσπάθειες ανάπτυξης πλατφόρμας που υποστήριζε την δημιουργία network abstractions ήταν η πλατφόρμα VeRTIGO [Dori12]. Η συγκεκριμένη πλατφόρμα βασίστηκε στην ύπαρξη του FlowVisor και στην δημιουργία ενός επιπέδου αφαίρεσης δικτύου πάνω από αυτόν, με την δυνατότητα παρουσίασης διαφορετικών όψεων της τοπολογίας σε διαφορετικούς OpenFlow controllers χρηστών. Οι VNE αλγόριθμοι που χρησιμοποιήθηκαν στην πλατφόρμα VeRTIGO [Rigg13] εφαρμόζονταν στο δίκτυο με τον καθορισμό του flowspace, δίχως να γίνεται μελέτη για το ποιοι μπορεί να είναι οι περιορισμοί όσον αφορά το πλήθος των αιτήσεων που μπορούσαν να γίνουν αποδεκτές και πιθανών περιορισμών κλίμακας της λύσεως (scalability issues) στο επίπεδο του FlowVisor. Στην εργασία μας μελετούμε τους πιθανούς περιορισμούς χρήσης proxy controllers όταν τα αιτήματα των χρηστών είναι χιλιάδες και οι δικτυακές τοπολογίες είναι ευρείας περιοχής (WAN).

Η χρήση πόρων του δικτύου από πολλαπλούς χρήστες, με εγκαθίδρυση διαφορετικών κανόνων ανά χρήστη σε ένα περιβάλλον κεντρικοποιημένου ελέγχου [Casa10], εισήγαγε την ιδέα των διαμοιραζόμενων ελεγκτών υποδομής (infrastructure controllers). Μια από τις πρώτες υλοποιήσεις ενός διαμοιραζόμενου OpenFlow controller αποτελεί ο FlowN [Drut13], ο οποίος επεκτείνει τον NOX controller [NOX] ώστε να δώσει την ικανότητα σε πολλαπλούς χρήστες να υλοποιούν τους δικούς τους αλγορίθμους προώθησης πακέτων μέσα στο ένα και μοναδικό σημείο ελέγχου του δικτύου, ο καθένας στο δικό του κομμάτι του flowspace. Η δική μας εργασία είναι

συμπληρωματική του FlowN καθώς εισάγει την ιδέα των γενικών πολιτικών τμηματοποίησης του flowspace που μπορεί να υλοποιηθεί μέσα σε ένα infrastructure controller.

Σημαντικός παράγοντας που επηρεάζει την ανθεκτικότητα και την αποκρισιμότητα ενός τμηματοποιημένου κεντρικοποιημένου επιπέδου ελέγχου (centralized sliced control-plane), είναι οι δικτυακές καθυστερήσεις που εισάγονται στην επικοινωνία, μεταξύ των OpenFlow switches και του OpenFlow proxy controller καθώς και στις καθυστερήσεις που εισάγονται μεταξύ του OpenFlow proxy controller και των controllers των χρηστών. Η επιλογή κατάλληλης θέσης στην τοπολογία του δικτύου για την τοποθέτηση των οντοτήτων που εκτελούν λειτουργίες του επιπέδου ελέγχου επηρεάζει την συνολική λειτουργία και απόδοση του δικτύου καθώς και την ανθεκτικότητα του. Το συγκεκριμένο πρόβλημα έχει μελετηθεί στην βιβλιογραφία [Hell12]. Η μελέτη της μείωσης μέσης καθυστέρησης καθώς και της μείωσης χείριστης καθυστέρησης επικοινωνίας μεταξύ του proxy controller και των OpenFlow switches έχει δείξει ότι ακόμα και σε δίκτυα ευρείας περιοχής, ένας μικρός αριθμός από καλά τοποθετημένους proxy controllers μπορεί να λειτουργήσει ικανοποιητικά, αφού εισάγει ανεκτές καθυστερήσεις της τάξης των μερικών δεκάδων milliseconds.

Οι καθυστερήσεις μετάδοσης που μελετώνται στο [Hell12] δρουν αθροιστικά με τις καθυστερήσεις επεξεργασίας των μηνυμάτων ελέγχου στο επίπεδο τμηματοποίησης ελέγχου, είτε στην περίπτωση χρήσης proxy controller, είτε στην περίπτωση χρήσης network hypervisor. Συνεπώς, η μέτρηση των καθυστερήσεων επεξεργασίας μηνυμάτων ελέγχου στο επίπεδο τμηματοποίησης ελέγχου (flowspace slinging), λόγω της εφαρμογής των πολικών τμηματοποίησης που μελετούμε παρακάτω, αποκτά νόημα για την εκτίμηση της ανθεκτικότητας και της αποκρισιμότητας στην λογική που επιβάλουν οι OpenFlow controllers των χρηστών.

6.3 Αρχές Σχεδίασης και Αρχιτεκτονικές

Σε αυτή την παράγραφο θα αναφερθούμε στις δυο αρχιτεκτονικές που μπορούν να διαμοιράσουν πόρους σε περιβάλλοντα SDN πολλαπλών χρηστών. Η πρώτη αρχιτεκτονική (proxy-controller architecture) κάνει χρήση ενός ενδιάμεσου OpenFlow

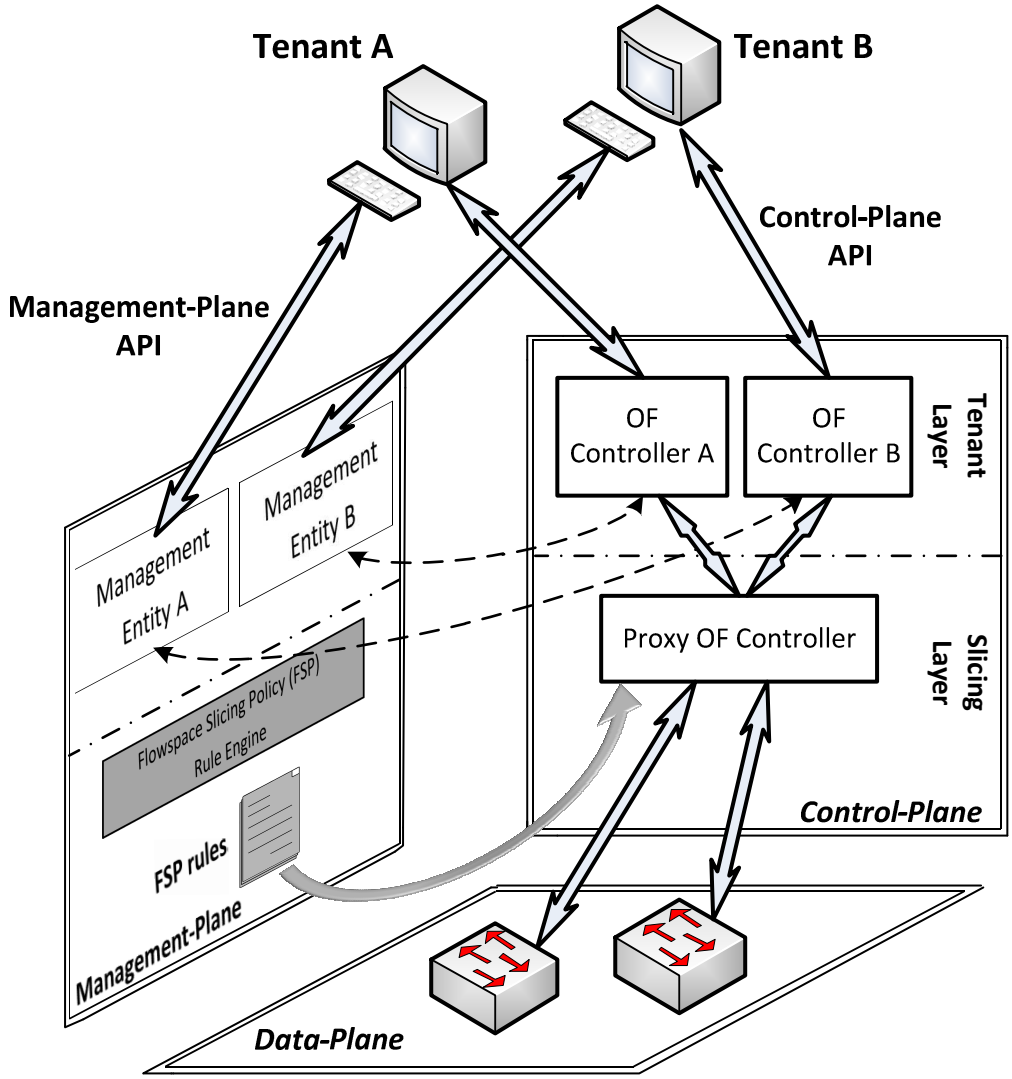
controller, ο οποίος επιτρέπει στους χρήστες να μοιράζονται τους πόρους του επιπέδου ελέγχου (όσον αφορά τις ροές δικτύου). Η συγκεκριμένη αρχιτεκτονική επιτρέπει σε κάθε χρήστη να αναπτύσσει την δική του λογική προώθησης πακέτων που ανήκουν στις προκαθορισμένες ροές που του έχουν αποδοθεί.

Η δεύτερη αρχιτεκτονική κάνει χρήση μιας οντότητας που ονομάζεται Network Hypervisor. Η συγκεκριμένη οντότητα επιτρέπει την δημιουργία ενός στρώματος αφαίρεσης δικτύου παρουσιάζοντας στους χρήστες των εικονικών δικτύων τοπολογίες διαφορετικές από την πραγματική τοπολογία υποδομής.

6.3.1 Αρχιτεκτονική Proxy Controller

Η απόδοση του flowspace σε δίκτυα οριζόμενα από λογισμικό μπορεί να υλοποιηθεί χρησιμοποιώντας ένα ενδιάμεσο επίπεδο διαμοιρασμού (slicing layer of the control-plane) στο επίπεδο ελέγχου, όπως φαίνεται στο **Σχήμα 17**. Κάθε μήνυμα ελέγχου του πρωτοκόλλου OpenFlow περνά αυτό το ενδιάμεσο επίπεδο και προωθείται στον OpenFlow controller του χρήστη, σύμφωνα με την πολιτική που εφαρμόζει ο OpenFlow proxy controller. Το επίπεδο που παίρνονται οι αποφάσεις για την προώθηση των πακέτων μέσα στο δίκτυο είναι αυτό στο οποίο ανήκουν όλοι οι OpenFlow controllers (OFC) των χρηστών (tenant layer of the control-plane). Κατά αντίστοιχο τρόπο τα μηνύματα, που δημιουργούνται από τους OpenFlow controllers των χρηστών και έχουν τελικό αποδέκτη τα OpenFlow switches στο επίπεδο προώθησης δεδομένων (data-plane), αντιπαραβάλλονται με την εφαρμοζόμενη πολιτική στον OpenFlow proxy controller.

Η πολιτική διαμοιρασμού του flowspace των χρηστών δομείται βάσει μιας μεθόδου (αλγορίθμου) που εγγυάται τη δημιουργία μη επικαλυπτόμενων περιοχών και υλοποιείται με τους κανόνες εφαρμογής στο slicing layer. Το **Σχήμα 17** υιοθετεί την αρχή διαχωρισμού λειτουργιών μεταξύ data-plane και control-plane καθώς και τη λειτουργία ενός κάθετου επιπέδου διαχείρισης, βασιζόμενο σε πολιτική (vertical Policy Based Network Management plane - PBNM), όπως αυτό ορίζεται από μοντέλα αναφοράς διαφόρων οργανισμών προτυποποίησης [RFC3198] [RFC3460] [DMTF].



Σχήμα 17: Αρχιτεκτονική διαμοιρασμού του επιπέδου ελέγχου με χρήση Proxy Controller

Οι αναλογίες που δημιουργούνται μεταξύ του μοντέλου PBNM και της αρχιτεκτονικής διαμοιρασμού του επιπέδου ελέγχου με χρήση Proxy controller που παρουσιάζουμε υπαγορεύουν την χρήση της προτεινόμενης μηχανής FSP ως το σημείο απόφασης της πολιτικής (Policy Decision Point – PDP). Η μηχανή FSP είναι υπεύθυνη για την μετάφραση των πολιτικών του διαχειριστή υποδομής και των αιτήσεων των χρηστών σε flowspace κανόνες. Οι flowspace κανόνες εφαρμόζονται στο σημείο εφαρμογής πολιτικής (Policy Enforcement Point – PEP) που είναι ο OpenFlow proxy controller [PBNM] [NMKA].

6.3.2 Αρχιτεκτονική Network Hypervisor

Η δεύτερη αρχιτεκτονική διαμοιρασμού του επιπέδου ελέγχου σε δίκτυα οριζόμενα από λογισμικό βασίζεται στην ύπαρξη ενός Network Hypervisor, όπως φαίνεται στο **Σχήμα 18**. Η συγκεκριμένη αρχιτεκτονική αποτελείται από τρία (3) διακριτά στρώματα στο επίπεδο ελέγχου:

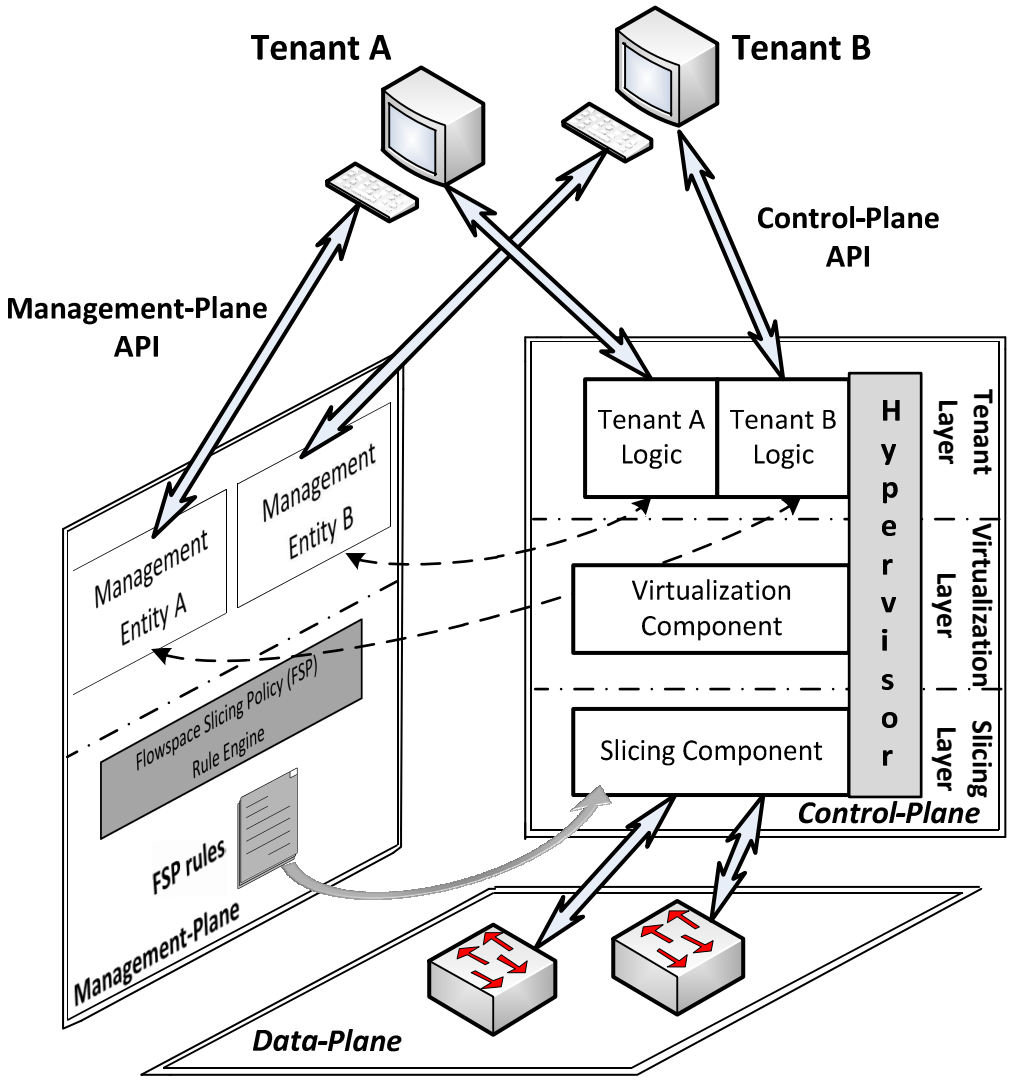
(i) Το στρώμα διαμοιρασμού του επιπέδου ελέγχου (slicing layer of the control-plane) είναι υπεύθυνο για την εφαρμογή των κανόνων διαμοιρασμού του flowspace.

(ii) Το στρώμα εικονικοποίησης (virtualization layer of the control-plane) είναι υπεύθυνο για την απόδοση του ελέγχου των πόρων της υποδομής στους χρήστες και δύναται να περιλαμβάνει λειτουργίες αφαίρεσης δικτύου. Ο χρήστης, μέσω της συγκεκριμένης διεπαφής, αντιλαμβάνεται την ύπαρξη ενός υποσυνόλου των πόρων (network discovery view per tenant), καθώς και να δημιουργεί εικονικές τοπολογίες που δεν αντιστοιχούν ευθέως με την τοπολογία υποδομής μέσω των λειτουργιών path migration και path splitting. Το στρώμα εικονικοποίησης μπορεί να λείπει από ορισμένες υλοποιήσεις που δεν διαθέτουν δυνατότητες εικονικοποίησης και αφαίρεσης σε επίπεδο τοπολογίας (full virtualization).

(iii)Στο επίπεδο ελέγχου των χρηστών (tenant layer of the control plane) εφαρμόζεται η εκάστοτε πολιτική προώθησης δεδομένων που θέλει να επιβάλλει ο κάθε χρήστης στο κομμάτι εκείνο του flowspace που του αντιστοιχεί. Η λογική προώθησης που επιβάλλει ο χρήστης μπορεί να είναι προκαθορισμένη και να προσφέρεται ως υπηρεσία από τον πάροχο υποδομής στο συγκεκριμένο επίπεδο λειτουργίας ή να είναι αρμοδιότητα του χρήστη να την αναπτύξει μόνος του.

Για λόγους κλιμακοθετησιμότητας (scalability) η αρχιτεκτονική του Network Hypervisor μπορεί να συμπεριλάβει ένα Proxy Controller με δυνατότητες αφαίρεσης δικτύου. Σε αυτή την περίπτωση η λογική προώθησης από πλευράς χρηστών

υλοποιείται από διακριτούς controllers που αλληλεπιδρούν με τον Proxy controller αυξημένων δυνατοτήτων μέσω του πρωτοκόλλου OpenFlow. Σε κάθε περίπτωση συνεχίζει να ισχύει η αρχή των διακριτών επιπέδων data-plane, control-plane και management-plane.



Σχήμα 18: Αρχιτεκτονική διαμοιρασμού του επιπέδου ελέγχου με χρήση Network Hypervisor

6.4 Μέθοδοι Διαμοιρασμού του Flowspace

Η εξυπηρέτηση πολλαπλών χρηστών σε μια δικτυακή τοπολογία που τους έχει αποδοθεί μέρος του ελέγχου όσον αφορά την προώθηση δεδομένων απαιτεί την ταξινόμηση των πακέτων σε flows, μέσω κάποιου πεδίου διαχωρισμού των επικεφαλίδων που έχουν όλα τα πακέτα (packet ID). Με αυτό τον τρόπο δίνεται η δυνατότητα ομαδοποίησης των πακέτων σε ροές, που με τη σειρά τους μπορεί αντιστοιχιστούν σε κάποιο χρήστη, στον οποίο ανατίθεται και ο έλεγχός τους.

Σε αυτό το πλαίσιο ένας OpenFlow controller, είτε είναι Proxy Controller, είτε είναι Network Hypervisor μπορεί να αποδίδει τον έλεγχο των ροών βάσει κάποιου πεδίου των επικεφαλίδων. Ο συγκεκριμένος μηχανισμός αντιστοιχεί μια τιμή ενός πεδίου επικεφαλίδας σε ένα slice της υποδομής. Το OpenFlow protocol μπορεί να χειρίζεται επικεφαλίδες από το Layer 2 έως το Layer 4 της στοίβας πρωτοκόλλων και κατά συνέπεια μπορεί να χρησιμοποιηθεί οποιαδήποτε συμβατή με το OpenFlow επικεφαλίδα ως λογικός διαχωριστής (βλέπε 0 για συμβατές επικεφαλίδες). *Σε περίπτωση που επιλεγεί κάποιο πεδίο διαχωρισμού από επικεφαλίδα του Layer 2 (π.χ. MPLS label ή VLAN ID) δίνεται η δυνατότητα στους χρήστες να ελέγχουν τα flows που τους αποδίδονται βάσει των υπολοίπων πεδίων επικεφαλίδων και άρα να έχουν το μέγιστο δυνατό έλεγχο πάνω στα πακέτα.*

Η χρήση ως packet ID ενός πεδίου, όπως είναι το VLAN ID και το MPLS label, είναι αναμενόμενη επιλογή διότι δεν δεσμεύει επικεφαλίδες που εξυπηρετούν άλλο σκοπό. Για παράδειγμα η επιλογή του 3-bit Priority code Point (PCP) από το IEEE 802.1Q πρωτόκολλο θα μπορούσε να αποτελέσει λογικό διαχωριστή, αλλά θα είχε το μειονέκτημα ότι ο μέγιστος αριθμός slices θα ήταν οχτώ και δεν θα επιτρεπόταν η παράλληλη χρήση του ως πεδίο απόδοσης προτεραιότητας μεταξύ των Ethernet πλαισίων (σκοπός του βάσει IEEE 802.1Q). Το MPLS label υποστηρίζεται από την έκδοση OpenFlow 1.1 και μεταγενέστερες, αλλά η αυξημένη πολυπλοκότητα υλοποίησης δεν έχει επιτρέψει μέχρι στιγμής την υποστήριξή του σε προϊόντα παραγωγής. Συνεπώς, ασφαλής λύση για τον διαχωρισμό αποτελεί μόνο το VLAN ID που γνωρίζει μια εκτενέστατη υποστήριξη και σε OpenFlow switches, αλλά και σε υπάρχοντα περιβάλλοντα δικτύωσης.

Το πρόβλημα με την χρήση του VLAN ID ως Packet ID είναι ότι μπορεί να λάβει 4096 διαφορετικές τιμές, αριθμός που σε μεγάλες υποδομές μπορεί να μην είναι ικανός να ικανοποιήσει το σύνολο των αιτήσεων για εικονικά δίκτυα. Για να αντιμετωπιστεί το συγκεκριμένο πρόβλημα κλιμακοθετησιμότητας, συμπεριλάβαμε ως παράμετρο λογικού διαχωρισμού των ροών το αναγνωριστικό του switch (switch ID) και το αναγνωριστικό της πόρτας του switch (switch port ID).

Η συγκεκριμένη ικανότητα διαχωρισμού θα μπορούσε να υποστηριχθεί από υπάρχοντες OpenFlow proxy controllers και δικτυακές υποδομές οριζόμενες από λογισμικό, δίχως την αλλαγή του OpenFlow πρωτοκόλλου. Οι proxy controllers θα μπορούσαν να παραμετροποιηθούν κατάλληλα από τις οντότητες συνάθροισης δικτυακών πόρων, που αποδίδουν στους χρήστες τους δικτυακούς πόρους (Aggregate Managers - AM). Χαρακτηριστικό παράδειγμα AM που αναπτύχθηκε για την υποστήριξη των πειραματικών υποδομών του GENI είναι ο FOAM [FOAM].

Συνεπώς , ένα εικονικό δίκτυο μπορεί να αντιστοιχιστεί με ένα μέρος του συνολικού flowspace της υποδομής με τους παρακάτω τρόπους:

(α) **domain-wide slicing**

Κάθε εικονικό δίκτυο είναι αυστηρά συνδεδεμένο με μια τιμή ενός Packet ID και το συγκεκριμένο Packet ID είναι δεσμευμένο σε όλη την δικτυακή υποδομή, π.χ. <VLAN ID>

(β) **switch-wide slicing**

Κάθε εικονικό δίκτυο αντιστοιχίζεται με πολλαπλούς λογικούς διαχωριστές που εκτός από το packet ID περιλαμβάνουν και τα switch IDs των συσκευών από τα οποία περνούν οι ροές του συγκεκριμένου χρήστη, π.χ. <VLAN ID, switch ID>

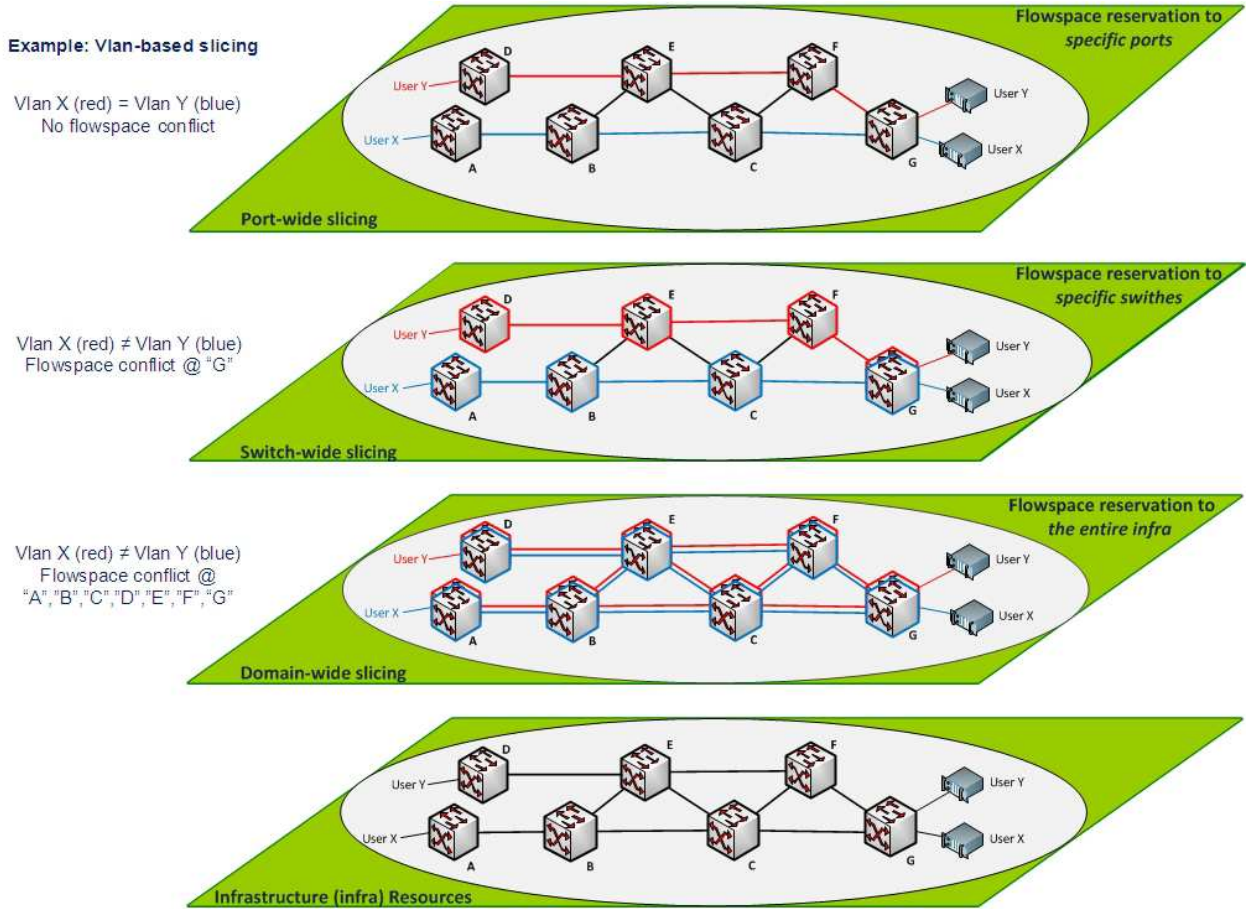
Για τον ορισμό του συνολικού flowspace χρήστη απαιτούνται πολλαπλές πλειάδες της ανωτέρω μορφής.

(γ) **port-wide slicing**

Στην τρίτη περίπτωση, κάθε εικονικό δίκτυο αντιστοιχίζεται με πολλαπλούς λογικούς διαχωριστές που εκτός από το packet ID περιλαμβάνουν και τα switch

IDs και τα port IDs από τις οποίες περνούν οι ροές του συγκεκριμένου χρήστη, π.χ. <VLAN ID, switch ID, switch port ID>

Για τον ορισμό του συνολικού flowspace που αντιστοιχεί με το flowspace απαιτούνται πολλαπλές πλειάδες της ανωτέρω μορφής.



Σχήμα 19: Μηχανισμοί ανάθεσης Flowspace σε SDN υποδομές με διαχωρισμό χρηστών (domain-wide, switch-wide, port-wide slicing) βάσει VLAN ID

Οι τρόποι της αποκλειστικής ανάθεσης VLAN IDs ανά χρήστη (α) σε όλη την τοπολογία, (β) σε ένα υποσύνολο των OpenFlow switches της τοπολογίας (switch-based slicing) και (γ) σε ένα υποσύνολο των διαθέσιμων πορτών των OpenFlow switches (port-based slicing), παρουσιάζονται στο **Σχήμα 19**.

Στην περίπτωση που ο χρήστης θέλει να του αποδοθεί ο έλεγχος συγκεκριμένου μέρους του flowspace σε μη επικαλυπτόμενες διαδρομές [Suur84] (disjoint paths) τίθεται το ερώτημα εύρεσης μη επικαλυπτόμενων διαδρομών, οι οποίες πρέπει να

έχουν αδέσμευτο το συγκεκριμένο μέρος του flowspace και εν συνεχεία η ανάθεση του flowspace στον χρήστη. Το κεντρικοποιημένο επίπεδο ελέγχου κάνει την εύρεση και απόδοση των διαδρομών πιο εύκολη, σε σχέση με δίκτυα που έχουν κατανεμημένο επίπεδο ελέγχου (distributed control-plane), αφού ο υπολογισμός τους και η δέσμευση των πόρων στο επίπεδο προώθησης δεδομένων (data-plane) γίνεται από το κεντρικό σημείο ελέγχου όλου του δικτύου, με πλήρη γνώση της τοπολογίας.

Αντιθέτως, δυσκολία παρουσιάζει το γεγονός ότι ο χρήστης δεν ζητά μόνο την εγκαθίδρυση μια διαδρομής ροής πακέτων στο επίπεδο προώθησης δεδομένων (data plane path reservation) αλλά την απόδοση σε αυτόν του συνολικού ελέγχου ενός μέρους του flowspace που αντιστοιχεί με αυτές (control plane flowspace delegation). Ζητά δηλαδή πόρους από το επίπεδο ελέγχου στο οποίο δεν είχε πρόσβαση σε κλασσικά δίκτυα (MPLS/GMPLS).

6.4.1 Πολιτικές Διαμοιρασμού του Flowspace

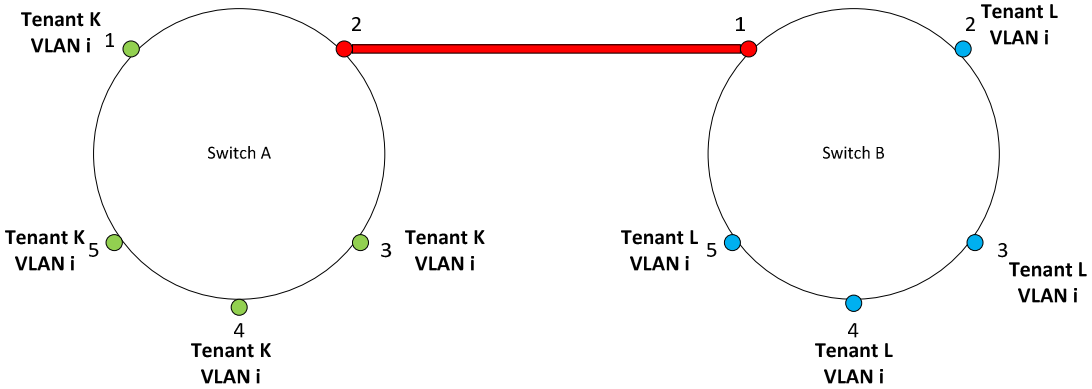
Ανεξαρτήτως της SDN αρχιτεκτονικής πολλαπλών χρηστών και της μεθόδου διαμοιρασμού του flowspace που θα υιοθετήσει ο πάροχος της υποδομής, θα πρέπει να υπάρχει ένα ισχυρός μηχανισμός απομόνωσης των χρηστών από την στιγμή που χρησιμοποιείται κοινή τοπολογία και κοινοί πόροι δικτύωσης [Sher10]. Στις δύο αρχιτεκτονικές που περιγράψαμε στην Παράγραφο 6.3 την ευθύνη για την απομόνωση των χρηστών, μέσω της ύπαρξης μη επικαλυπτόμενων κανόνων για το flowspace, την αναλαμβάνει το control-plane slicing layer. Παρακάτω αναλύουμε τις συνέπειες που έχει η εφαρμογή της απομόνωσης.

Στην περίπτωση του domain-wide slicing θα πρέπει να γίνεται έλεγχος των ήδη χρησιμοποιούμενων Packet IDs, ώστε να αποφεύγεται η επαναχρησιμοποίησή τους σε οποιοδήποτε κομμάτι της τοπολογίας υποδομής από δεύτερο χρήστη.

Η εξασφάλιση της απομόνωσης μεταξύ των χρηστών στην περίπτωση του switch-wide slicing είναι πολυπλοκότερη, καθώς ένα Packet ID μπορεί να χρησιμοποιείται από περισσότερα εικονικά δίκτυα σε διαφορετικά OpenFlow switches της υποδομής. Στη περίπτωση μάλιστα που κακόβουλοι χρήστες ελέγχουν κάποιο κομμάτι του flowspace σε ένα σύνολο από switches, ενδέχεται να προσπαθήσουν να αποστείλουν, μέσω των

ζεύξεων υποδομής, κίνηση από switch που ελέγχουν σε παρακείμενο switch που δεν έχουν τον έλεγχο του flowspace. Σε αυτή την περίπτωση τα πακέτα που θα εισέλθουν στο παρακείμενο switch, αν υπάρχουν προεγκατεστημένα flow entries που αντιστοιχούν με τους headers τους, θα προωθηθούν αναλόγως μέσα στο εικονικό δίκτυο άλλου χρήστη. Σε περίπτωση που δεν υπάρχει κάποιο flow entry, το OpenFlow switch θα αναγκαστεί να ρωτήσει τον OpenFlow controller του «ξένου» εικονικού δικτύου για το πώς θα πρέπει να προωθηθούν τα πακέτα. Το συγκεκριμένο ενδεχόμενο μπορεί να δημιουργήσει καταιγισμό μηνυμάτων στο control-plane (control-plane overloading) άλλου χρήστη, επηρεάζοντας πιθανόν συνολικά την λειτουργία του εικονικού του δικτύου.

Στο **Σχήμα 20** παρουσιάζεται ενδεχόμενη παραβίαση της απομόνωσης μεταξύ των *χρηστών K* και *L*, που ελέγχουν τα παρακείμενα switches A και B αντίστοιχα. Τα δυο switches είναι διασυνδεδεμένα μέσω της *πόρτας 2* (switch A) και της *πόρτας 1* (*switch B*) και στους δυο χρήστες τους έχει αποδοθεί το ίδιο packet ID (π.χ *VLAN ID i)*. Εάν ο *χρήστης K* επιλέξει την *πόρτα 2* ως πόρτα εξόδου στο *switch A* που ελέγχει, τα πακέτα θα προσεγγίσουν το *switch B* στην *πόρτα 1*. Τα συγκεκριμένα πακέτα μπορεί να δημιουργήσουν υπερφόρτωση του επιπέδου ελέγχου σε περίπτωση που έχουν επιλεγεί με τέτοιο τρόπο ώστε να δημιουργούν ερωτήματα στο κεντρικοποιημένο επίπεδου ελέγχου του παρακείμενου χρήστη ή ακόμα και στο control-plane slicing layer ολόκληρης της υποδομής.



Σχήμα 20: Εφαρμογή της απομόνωσης χρηστών στην περίπτωση της switch-wide μεθόδου

Συνεπώς, η *πόρτα 2* δεν θα πρέπει να χρησιμοποιείται σε συνδυασμό με το *VLAN ID i* ούτε από τον *χρήστη K*, ούτε από τον *χρήστη L*. Το απαιτούμενο flowspace όσον αφορά το *χρήστη K*, στο *switch A* θα πρέπει να ορισθεί από τους παρακάτω 4 κανόνες (1 κανόνας ανά πόρτα):

tenant K, priority=1, in/out port=1,datapath=switch A, VLAN=i

tenant K, priority=1, in/out port=3,datapath=switch A, VLAN=i

tenant K, priority=1, in/out port=4,datapath=switch A, VLAN=i

tenant K, priority=1, in/out port=5,datapath=switch A, VLAN=i

Αντίστοιχα, όταν χρησιμοποιείται η port-wide slicing μέθοδος, οι κανόνες για το flowspace του κάθε εικονικού δικτύου, θα πρέπει να είναι μη επικαλυπτόμενοι για κάθε πόρτα της τοπολογίας. Όπως παρουσιάζεται στο **Σχήμα 21**, ο *χρήστης K* έχει δεσμεύσει τις *πόρτες 1* και *5* από το *switch A*, ενώ ο *χρήστης L* έχει δεσμεύσει τις *πόρτες 2* και *3* από το ίδιο switch. Οι δυο χρήστες χρησιμοποιούν το *VLAN i* ως Packet ID και για το λόγο αυτό αν κάποιος από τους δυο επιλέξει ως πόρτα εξόδου μια πόρτα που ανήκει στο γειτονικό χρήστη θα έχουμε παραβίαση της απομόνωσης. Για να αποφύγουμε τέτοια περίπτωση θα πρέπει η εφαρμογή της πολιτικής της port-wide μεθόδου να δημιουργεί μη επικαλυπτόμενους κανόνες για το flowspace που να περιλαμβάνουν και τις πόρτες ανά χρήστη:

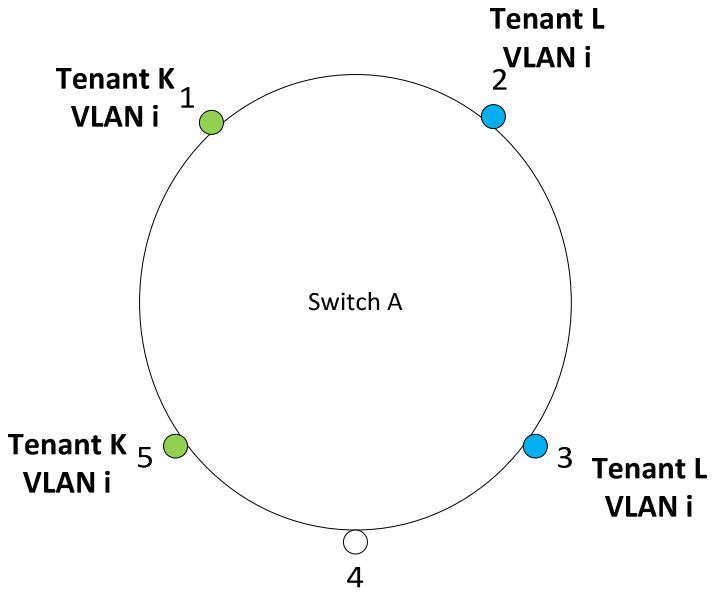
tenant K, priority=1, in/out port=1, datapath=switch A, VLAN=i

tenant K, priority=1, in/out port=5, datapath=switch A, VLAN=i

tenant L, priority=1, in/out port=2, datapath=switch A, VLAN=i

tenant L, Priority=1, in/out port=3, datapath=switch A, VLAN=i

Όπως παρατηρούμε οι δυο παραπάνω υλοποιήσεις των switch-wide και port-wide μεθόδων οδηγούν στη χρήση του ίδιου αριθμού κανόνων, αφού η περιγραφή του flowspace, για λόγους απομόνωσης των χρηστών, γίνεται ανά switch port.



Σχήμα 21: Εφαρμογή της απομόνωσης χρηστών στην περίπτωση της port-wide μεθόδου

6.4.2 Μείωση Κανόνων Flowspace στη Switch-wide Μέθοδο

Στην Παράγραφο 6.4.1 διαπιστώσαμε ότι ο αριθμός των κανόνων περιγραφής του flowspace δοθείσας εικονικής τοπολογίας με χρήση των μεθόδων switch-wide και port-wide είναι ίδιος. Για λόγους ικανοποιητικής απομόνωσης μεταξύ των εικονικών τοπολογιών απαιτείται ακόμα και στην switch-wide μέθοδο να γίνεται καθορισμός κανόνων για το flowspace ανά switch port. Παρακάτω παρουσιάζουμε ένα τρόπο μείωσης του συνολικού αριθμού των κανόνων στην περίπτωση που χρησιμοποιείται η switch-wide μέθοδος.

Για δοθέν packet ID (π.χ. VLAN ID) ένα σύνολο πορτών ενός switch παραχωρούνται (delegated/assigned ports) σε ένα συγκεκριμένο χρήστη. Πόρτες που διασυνδέουν switches (transit ports) της τοπολογίας δεν πρέπει να παραχωρούνται σε χρήστες (non-delegated ports), εκτός και αν τα δυο διασυνδεόμενα switches ανήκουν στον ίδιο χρήστη. Οι λόγοι έχουν αναφερθεί εκτενώς στην προηγούμενη Παράγραφο.

Όταν ο *αριθμός d των παραχωρηθέντων πορτών ανά switch* είναι μεγαλύτερος από τον *αριθμό n των μη παραχωρημένων πορτών* μπορούμε να μειώσουμε τον αριθμό των απαιτούμενων κανόνων με την παρακάτω μέθοδο:

(i) Χρησιμοποιούνται χαμηλής προτεραιότητας κανόνες με πεδία αναπλήρωσης (wildcards) για την απόδοση μέρους του flowspace στους χρήστες

(ii) Χρησιμοποιούνται υψηλής προτεραιότητας κανόνες απόρριψης πακέτων (drop rules) για τις transit ports, που διαμορφώνονται από ένα OpenFlow controller διαχειριστικών σκοπών. Ο συγκεκριμένος controller έχει την οπτική ολόκληρης της τοπολογίας υποδομής και εισάγει τους drop rules που απαιτούνται. *Οι κανόνες υψηλής προτεραιότητας αποτρέπουν τη μετάδοση κίνησης από εικονικό σε εικονικό δίκτυο (traffic injection) και διατηρούν το επίπεδο ελέγχου κάθε χρήστη απομονωμένο από ερωτήματα για την προώθηση πακέτων που δεν ανήκουν στο εικονικό τους δίκτυο.*

Θα χρησιμοποιήσουμε ως παράδειγμα τους κανόνες που απαιτούνται για το switch A της Παραγράφου 6.4.1 και την τοπολογία που παρουσιάσαμε στο **Σχήμα 20** για να δείξουμε ότι μπορεί να μειωθεί ο αριθμός των απαιτούμενων κανόνων βάσει της παραπάνω μεθοδολογίας (βλέπε **Πίνακας 6-1)**. Η ποσοστιαία μείωση κανόνων είναι μεγαλύτερη όσο μεγαλύτερος είναι ο αριθμός των πορτών που αποδίδονται στο χρήστη, αφήνοντας μικρότερο αριθμό μη αποδοθέντων πορτών.

Flowspace Reservation				
Tenant id	**Datapath id**	**Vlan tag id**	**Tenant ports**	
Tenant K	Switch A	Vlan i	Ports 1,3,4,5	
Tenant L	Switch B	Vlan i	Ports 2,3,4,5	
Flowspace Implementation				
Slice id	**Priority**	**Port**	**Datapath id**	**Vlan tag id**
Reserved slice 1	Priority 1	Port=2	Switch A	Vlan tag i
Tenant K slice	Priority 2	Port=*	Switch A	Vlan tag i
Reserved slice 1	Priority 1	Port=1	Switch B	Vlan tag i
Tenant L slice	Priority 2	Port=*	Switch B	Vlan tag i

Πίνακας 6-1: Μείωση των απαιτούμενων κανόνων flowspace όταν d > n με χρήση της switch-wide μεθόδου

Όταν ο αριθμός d των πορτών ανά switch είναι μικρότερος από τον αριθμό n του ίδιου switch, ο αριθμός των κανόνων που χρειάζονται για το συγκεκριμένο switch είναι ίσος με τον αριθμό των αποδιδομένων πορτών (**Πίνακας 6-2)**.

Flowspace Reservation				
Tenant id	**Datapath id**	**Vlan tag id**	**Tenant ports**	
Tenant K	Switch A	Vlan i	Ports 1,4	
Tenant L	Switch B	Vlan i	Ports 2,4	
Flowspace Implementation				
Slice id	**Priority**	**Port**	**Datapath id**	**Vlan tag id**
Tenant K slice	Priority 1	Port=1	Switch A	Vlan tag i
Tenant K slice	Priority 1	Port=4	Switch A	Vlan tag i
Tenant L slice	Priority 1	Port=2	Switch B	Vlan tag i
Tenant L slice	Priority 1	Port=4	Switch B	Vlan tag i

Πίνακας 6-2: Αριθμός των απαιτούμενων κανόνων flowspace όταν d > n με χρήση της switch-wide μεθόδου

6.5 Πειραματικές Μετρήσεις

Στην Παράγραφο 6.5.1 παρουσιάζουμε την μεθοδολογία πειραματικής αξιολόγησης για τις τρεις μεθόδους διαμοιρασμού του επιπέδου δικτυακού ελέγχου (domain-wide, switch-wide, port-wide). Υλοποιούμε μια μηχανή διαμοιρασμού του flowspace (Flowspace slicing policy -FSP) που εφαρμόζει τις τρεις μεθόδους [FSPengine].

Στην Παράγραφο 6.5.2 παρουσιάζουμε τα αποτελέσματα των πειραματικών μετρήσεων που έγιναν με χρήση της μηχανής FSP καθώς και του FlowVisor που ανέλαβε το έργο της εφαρμογής της πολιτικής απομόνωσης των εικονικών δικτύων.

6.5.1 Μεθοδολογία Πειραματικής Αξιολόγησης

Η πειραματική αξιολόγηση της απόδοσης των τριών διαφορετικών μεθόδων διαμοιρασμού του επιπέδου ελέγχου έγινε με την βοήθεια της FSP. Η μηχανή FSP λαμβάνει ως είσοδο τις αιτήσεις για τη δημιουργία εικονικών δικτύων/τοπολογιών των χρηστών και την τοπολογία υποδομής πάνω στην οποία θα πρέπει να δημιουργηθούν τα εικονικά δίκτυα και δημιουργεί το σύνολο των απαιτούμενων κανόνων που καθορίζουν το flowspace.

Τα σενάρια που μελετήθηκαν περιελάμβαναν την δημιουργία των κανόνων του flowspace και την εισαγωγή τους στον FlowVisor για να γίνει εκτίμηση της χρονικής επιβάρυνσης που εισάγει καθώς και της κατανάλωσης μνήμης συστήματος που απαιτεί. Οι δικτυακές τοπολογίες υποδομής που χρησιμοποιήθηκαν περιελάμβαναν το Internet2/OS3E [IN2topo] και το GÉANT [GEANTtopo] καθώς και άλλες τοπολογίες ακαδημαϊκών κυρίως δικτύων που ήταν διαθέσιμες στο [INtopo]. Τα αιτήματα των χρηστών για δημιουργία εικονικών δικτύων που θα επέτρεπαν τη διασύνδεση σημείων της τοπολογίας υποδομής ικανοποιήθηκαν με βάση αλγορίθμους εύρεσης της μικρότερης διαδρομής (shortest-path algorithms), με κριτήριο την καθυστέρηση διάδοσης (propagation delay).

Οι χρήστες είχαν τη δυνατότητα να επιλέξουν συγκεκριμένη τιμή του Packet ID σε ορισμένα πειραματικά σενάρια (**bound requests**), ενώ σε κάποια άλλα άφηναν τη συγκεκριμένη επιλογή να γίνει από το επίπεδο διαχείρισης της υποδομής (**unbound requests**). Η επιλογή του Packet ID από τον χρήστη είναι ένα ενδεχόμενο που μελετήθηκε, καθώς μπορεί να επιφέρει απλοποίηση διαχειριστικών λειτουργών στο επίπεδο του εικονικού δικτύου. Η έξοδος της FSP μηχανής (κανόνες flowspace) ανά μέθοδο ήταν σύμφωνη με όσα έχουν ήδη περιγραφεί στις Παραγράφους 6.4.1 και 6.4.2.

Στην πειραματική διάταξη που αναπτύξαμε, κάθε αίτηση χρήστη αντιστοιχιζόταν με μια τυχαία εικονική τοπολογία. Οι τρεις κατηγορίες τοπολογιών που ήταν διαθέσιμες και βάσει των οποίων δημιουργούνταν όλα τα αιτήματα εικονικών δικτύων

περιελάμβαναν: (α) τοπολογίες αστέρα, (β) δισημειακές διαδρομές (point-to-point paths) και μη επικαλυπτόμενες διαδρομές (disjoint paths).

Η κατηγορία τοπολογιών αστέρα επιλέχθηκε καθώς αποτελούν συνηθισμένη απαίτηση των παρόχων περιεχομένου (content providers, π.χ. NetFlix, Hulu) για τη δημιουργία Content Delivery Networks (CDNs). Η κατηγορία των δισημειακών διαδρομών επιλέχθηκε γιατί καλύπτει ένα σύνολο αναγκών των εναλλακτικών παρόχων (alternative providers) για τη διασύνδεση σημείων παρουσίας τους και τέλος οι μη επικαλυπτόμενες διαδρομές είναι μια ειδική κατηγορία αιτημάτων με στόχο την αύξηση της διαθεσιμότητας συνδέσεων κρίσιμης σημασίας. Ο συνολικός αριθμός αιτημάτων για μη επικαλυπτόμενες διαδρομές που χρησιμοποιήθηκαν καθορίστηκε από το μέγιστο αριθμό των υπαρκτών μη επικαλυπτόμενων διαδρομών ανά τοπολογία υποδομής, που ήταν και το άνω όριο δυνατών διαδρομών. Ο αλγόριθμος που χρησιμοποιήθηκε για την αναζήτηση των διαδρομών στις τοπολογίες υποδομής περιγράφεται στο [Bhan97].

Τα διαδοχικά πειράματα που εκτελέσαμε αφορούσαν έναν αριθμό αιτημάτων για εικονικά δίκτυα που κυμαινόταν από 1000 έως 16000. Το ποσοστό με το οποίο συμμετείχαν οι κατηγορίες τοπολογιών που περιγράψαμε παραπάνω ανά περίπτωση μπορεί να συνοψιστεί σε τέσσερα (4) σενάρια μιγμάτων αιτημάτων:

(i) Μίγμα 1 – Mix 1

Το 49% των αιτημάτων αφορούσαν τοπολογίες αστέρα και το 49% των αιτημάτων δισημειακές διαδρομές. Το υπόλοιπο 2% αφορούσε μη επικαλυπτόμενες διαδρομές. Το σύνολό τους ήταν bound requests.

(ii) Μίγμα 2 – Mix 2

Το 69% των αιτημάτων αφορούσαν τοπολογίες αστέρα και το 29% των αιτημάτων δισημειακές διαδρομές. Το υπόλοιπο 2% αφορούσε μη επικαλυπτόμενες διαδρομές. Το σύνολό τους ήταν bound requests.

(iii) Μίγμα 3 – Mix 3

ΤΟ 69% των αιτημάτων αφορούσαν τοπολογίες αστέρα και το 29% των αιτημάτων δισημειακές διαδρομές. Το υπόλοιπο 2% αφορούσε μη επικαλυπτόμενες διαδρομές. Το 20% των αιτημάτων από κάθε κατηγορία ήταν bound requests.

(iv) Μίγμα 4 – Mix 4

Το 69% των αιτημάτων αφορούσαν τοπολογίες αστέρα και το 29% των αιτημάτων δισημειακές διαδρομές. Το υπόλοιπο 2% αφορούσε μη επικαλυπτόμενες διαδρομές. Το σύνολο των αιτημάτων ήταν unbound requests.

Η πολιτική αποδοχής των αιτημάτων από την FSP μηχανή διέφερε ανάλογα με την μέθοδο διαμοιρασμού που είχε επιλεγεί για δοκιμή και περιγράφεται παρακάτω:

- **Domain-wide slicing**

  - Ένα bound request γίνεται αποδεκτό όταν η τιμή του Packet ID που περιλαμβάνεται στο αίτημα είναι διαθέσιμη στο σύνολο της τοπολογίας υποδομής.

  - Ένα unbound request γίνεται αποδεκτό όταν υπάρχει τουλάχιστον μια τιμή του Packet ID διαθέσιμη στο σύνολο της τοπολογίας υποδομής.

- **Switch-wide slicing**

  - Ένα bound request γίνεται αποδεκτό όταν η τιμή του Packet ID του αιτήματος είναι διαθέσιμη σε κάθε switch της τοπολογίας υποδομής που περιλαμβάνει το εικονικό δίκτυο.

  - Ένα unbound request γίνεται αποδεκτό όταν υπάρχει τουλάχιστον μια τιμή του Packet ID διαθέσιμη σε κάθε switch της τοπολογίας υποδομής που περιλαμβάνει το εικονικό δίκτυο.

- **Port-wide slicing**

  - Ένα bound request γίνεται αποδεκτό όταν η τιμή του Packet ID του αιτήματος είναι διαθέσιμη σε κάθε πόρτα των switches της τοπολογίας υποδομής που περιλαμβάνει το εικονικό δίκτυο.

  - Ένα unbound request γίνεται αποδεκτό όταν υπάρχει τουλάχιστον μια τιμή του Packet ID διαθέσιμη σε κάθε πόρτα των switches της τοπολογίας υποδομής που περιλαμβάνει το εικονικό δίκτυο.

Τέλος καθορίζουμε τα μεγέθη σύγκρισης (metrics) που χρησιμοποιήσαμε για την αξιολόγηση των μεθόδων διαμοιρασμού του flowspace ως ακολούθως:

(i) Λόγος αποδοχής αιτημάτων - Acceptance ratio

Ο λόγος των συνολικών αιτημάτων που εξυπηρετήθηκαν από την FSP μηχανή προς το συνολικό αριθμό αιτημάτων που δέχθηκε. Το συγκεκριμένο μέγεθος

αποτελεί μια ένδειξη της συνολικής ευχαρίστησης/ικανοποίησης που δημιουργεί η πολιτική διαμοιρασμού του flowspace στους χρήστες**.**

(ii) Αριθμός των απαιτούμενων flowspace κανόνων – Number of required flowspace rules

Ο αριθμός των κανόνων που απαιτούνται για το διαμοιρασμό του επιπέδου ελέγχου σε πολλαπλούς χρήστες και την απομόνωση των εικονικών δικτύων των χρηστών στο επίπεδο ελέγχου και στο επίπεδο προώθησης δεδομένων.

(iii) Χρονική επιβάρυνση proxy controller – Proxy controller time overhead

Η χρονική καθυστέρηση που εισάγεται στο επίπεδο ελέγχου μεταξύ των OpenFlow switches και των OpenFlow controllers των χρηστών λόγω του επιπέδου διαμοιρασμού ελέγχου του proxy controller.

(iv) Κατανάλωση μνήμης proxy controller – Proxy controller memory consumption

Η απαιτούμενη μνήμη στον εξυπηρετητή που υλοποιεί τις λειτουργίες του proxy controller για τον χειρισμό των κανόνων του flowspace.

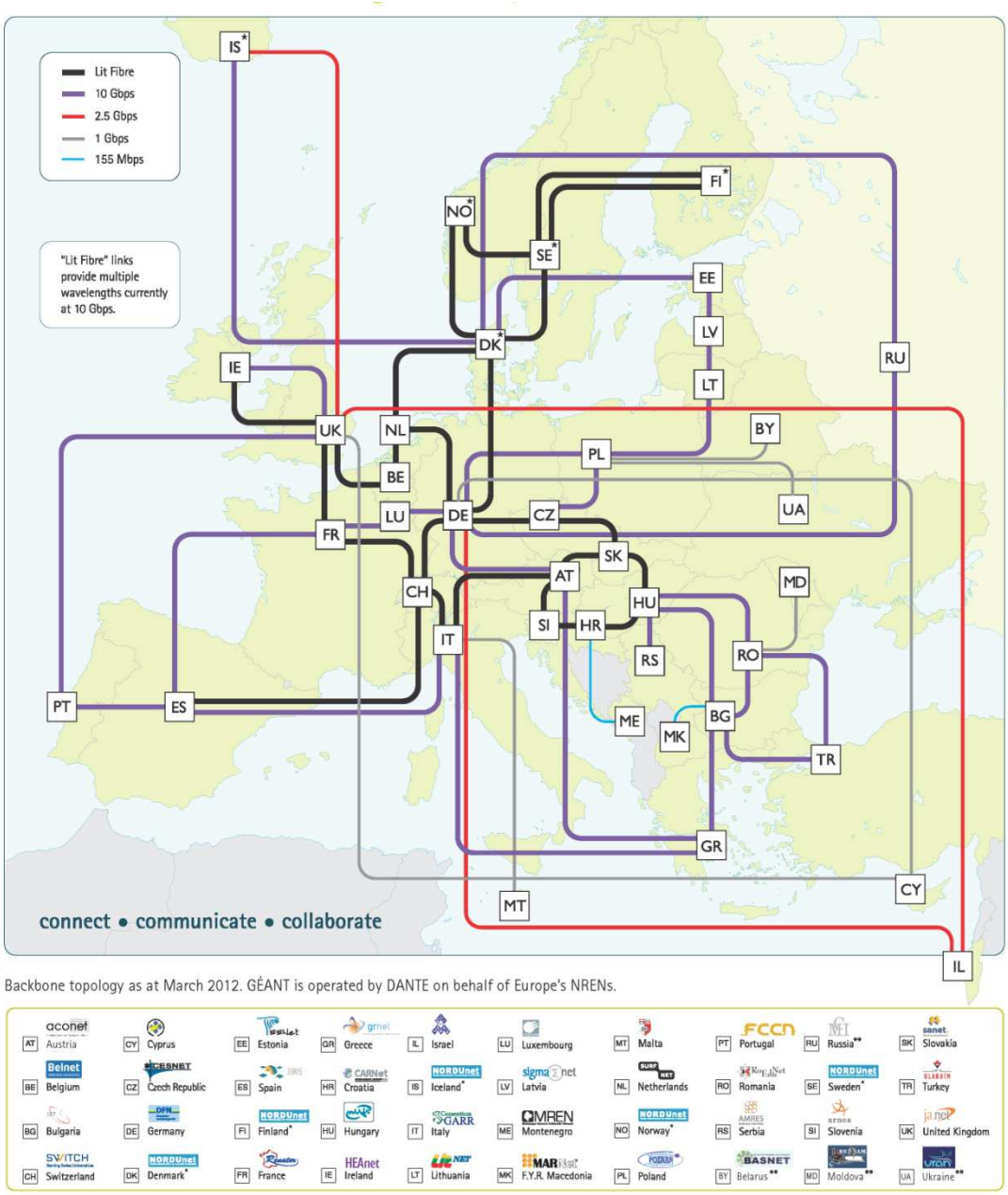
6.5.2 Αποτελέσματα Πειραματικών Μετρήσεων

Για την αξιολόγηση των μεθόδων διαμοιρασμού του flowspace χρησιμοποιούμε εκτενώς ως τοπολογίες υποδομής αυτές του GÉANT [GEANTtopo] και του Internet2/OS3E [IN2topo], όπως παρουσιάζονται στο **Σχήμα 22** και **Σχήμα 23**. Επίσης χρησιμοποιούμε τις τοπολογίες των παρακάτω δικτύων, όπως αυτές έχουν καταγραφεί στο [INtopo]:

Μικρές τοπολογίες (6 – 24 κόμβοι): Getnet (6 nodes), Ibm (17 nodes), Rediris (18 nodes), Belnet 2006 (22 nodes), Psinet (23 nodes)

Μεσαίες τοπολογίες (25 – 44 κόμβοι): Bbnplanet (26 nodes), CrlNetworkServices (32 nodes), Internet2/OS3E (34 κόμβοι), GÉANT (39 κόμβοι)

Μεγάλες τοπολογίες (>45 κόμβοι): Uunet (48 nodes), Ulaknet (81 nodes)

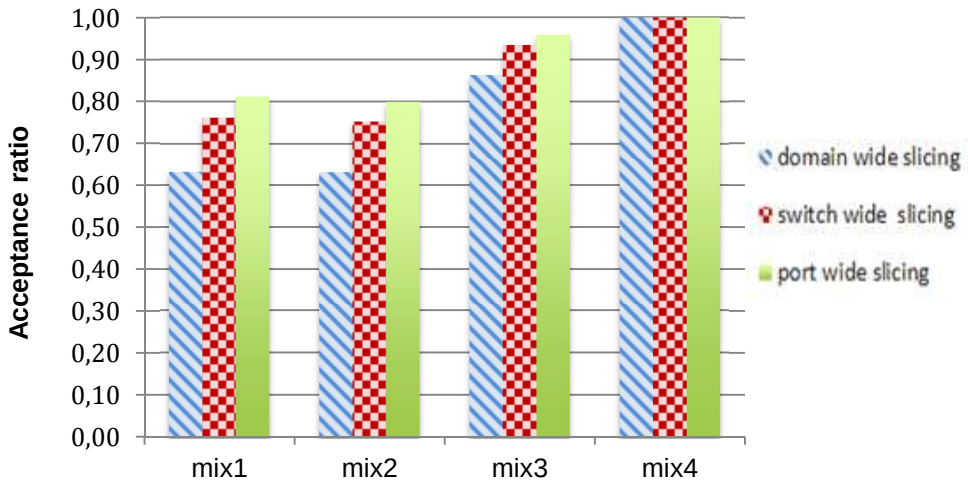


Backbone topology as at March 2012. GÉANT is operated by DANTE on behalf of Europe's NRENs.

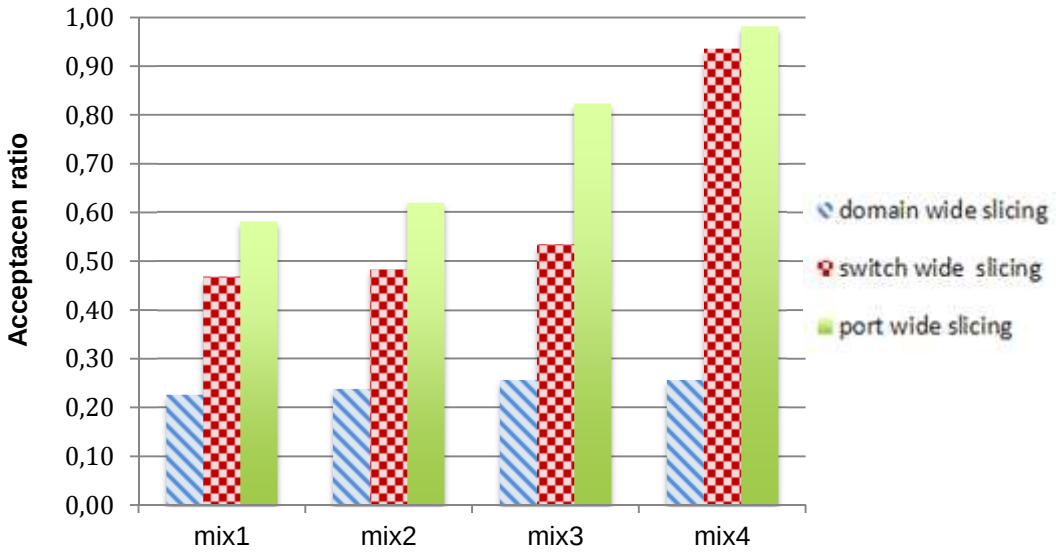
Σχήμα 22: Τοπολογία GÉANT (έτος 2012) πειραματικών μετρήσεων

Σχήμα 23: Τοπολογία Internet2/OS3E (έτος 2013) πειραματικών μετρήσεων

Ο λόγος αποδοχής τυχαίων αιτημάτων χρηστών για εικονικά δίκτυα πάνω στην τοπολογία υποδομής του Internet2/OS3E, συναρτήσει του μίγματος (mix1, mix2, mix3, mix4) και της μεθόδου (μπλε για domain-wide, κόκκινο για switch-wide και πράσινο για port-wide), για 4.000 και 16.000 αιτήματα φαίνεται στο **Σχήμα 24** και στο **Σχήμα 25** αντίστοιχα**.**

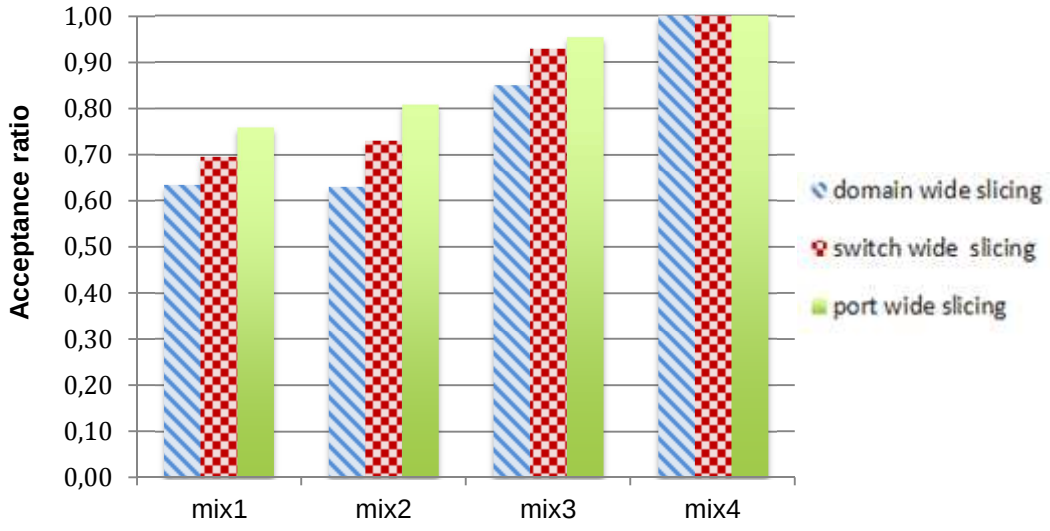


Σχήμα 24: Λόγος αποδοχής αιτημάτων χρηστών για εικονικά δίκτυα πάνω στο Internet2/OS3E για 4.000 αιτήματα

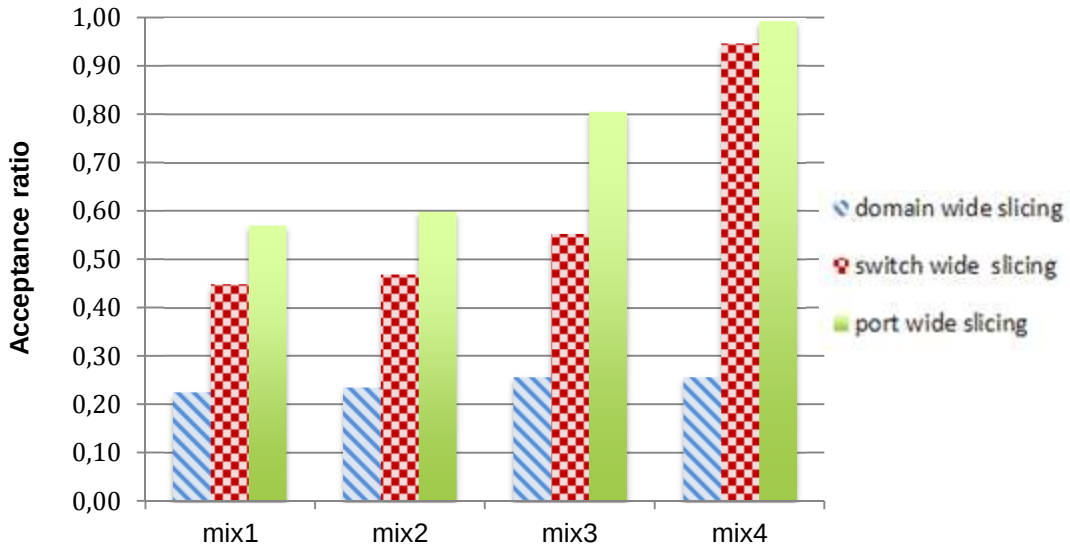


Σχήμα 25: Λόγος αποδοχής αιτημάτων χρηστών για εικονικά δίκτυα πάνω στο Internet2/OS3E για 16.000 αιτήματα

Ο λόγος αποδοχής τυχαίων αιτημάτων χρηστών για εικονικά δίκτυα πάνω στην τοπολογία υποδομής του GÉANT, συναρτήσει του μίγματος (mix1, mix2, mix3, mix4) και της μεθόδου (μπλε για domain-wide, κόκκινο για switch-wide και πράσινο για port-wide), για 4.000 και 16.000 αιτήματα φαίνεται στο **Σχήμα 26** και στο **Σχήμα 27** αντίστοιχα.



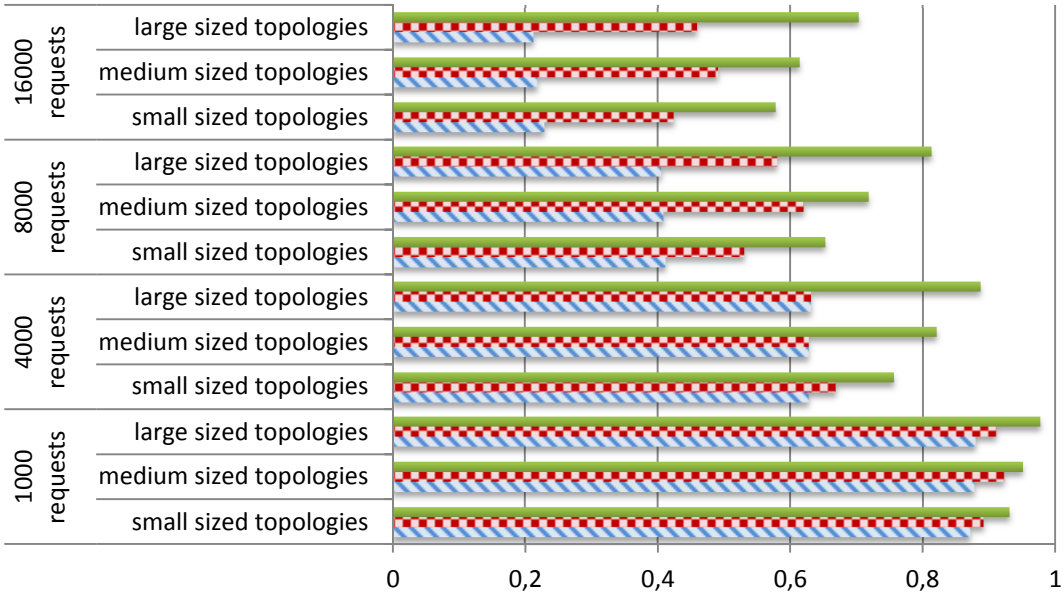
Σχήμα 26: Λόγος αποδοχής αιτημάτων χρηστών για εικονικά δίκτυα πάνω στο GÉANT για 4.000 αιτήματα



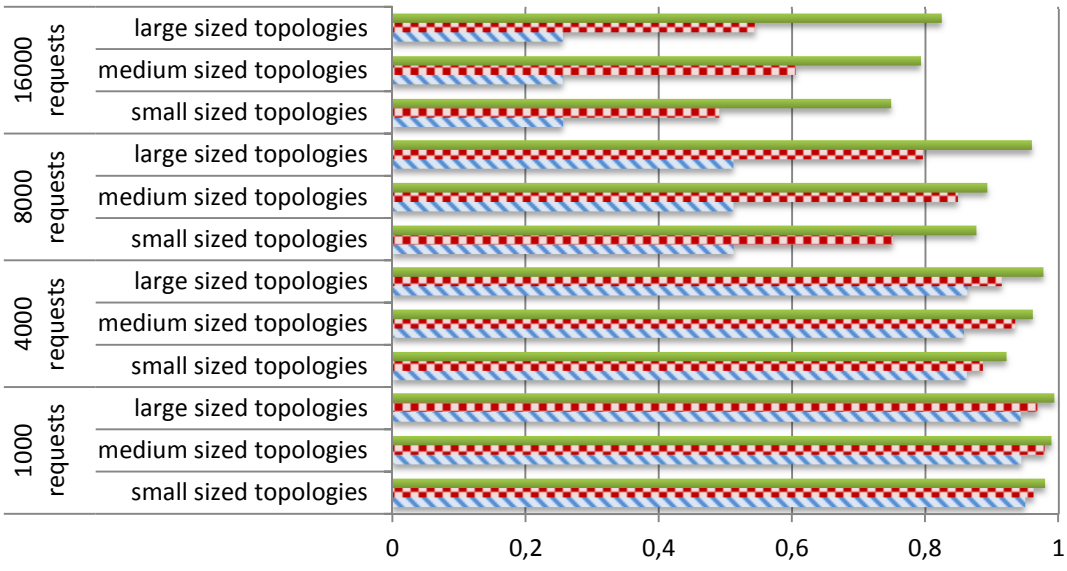
Σχήμα 27: Λόγος αποδοχής αιτημάτων χρηστών για εικονικά δίκτυα πάνω στο GÉANT για 16.000 αιτήματα

Από τα παραπάνω σχήματα και συγκρίνοντας τα αποτελέσματα για 4.000 και 16.000 αιτήσεις ανά τοπολογία, συμπεραίνουμε ότι όσο μεγαλύτερος είναι ο αριθμός των αιτήσεων που γίνονται για εγκαθίδρυση εικονικών δικτύων τόσο μειώνεται το ποσοστό αποδοχής αιτημάτων. Επίσης, ανεξαρτήτως του μίγματος των αιτήσεων που χρησιμοποιήθηκε για τις πειραματικές μετρήσεις, βλέπουμε ότι η μέθοδος port-wide slicing έχει την καλύτερη επίδοση και η domain-wide μέθοδος τη χειρότερη.

Σε περιπτώσεις που τα αιτήματα είναι πολύ λίγα σε σχέση με τον αριθμό των διαθέσιμων Packet IDs και η επιλογή των Packet IDs αφήνεται στον διαχειριστή της υποδομής που μπορεί να κάνει την βέλτιστη κατανομή τους οι τρεις μέθοδοι τείνουν να έχουν την ίδια επίδοση (mix4, 4.000 αιτήματα). Αντίθετα, όσο περισσότερα είναι τα αιτήματα σε σχέση με τον αριθμό των διαθέσιμων Packet IDs και η επιλογή των Packet IDs γίνεται από τους ίδιους τους χρήστες (mix1, 16.000) φαίνεται σαφώς ότι υπερτερεί η port-wide slicing μέθοδος. Σε αυτές μάλιστα τις περιπτώσεις η domain-wide μέθοδος δεν έχει καθόλου καλή επίδοση, με αποτέλεσμα να μπορεί να καλύψει μόνο το 20% των αιτημάτων, σε σχέση με το περίπου 60% των αιτημάτων που καλύπτει η port-wide slicing μέθοδος. Τέλος, η port-wide μέθοδος πετυχαίνει την καλύτερη επίδοση, συγκριτικά με τις άλλες δυο, όταν τα αιτήματα είναι πολλά και αφήνεται η επιλογή των Packet IDs στην υποδομή (mix4, 16.000).



Σχήμα 28: Λόγος αποδοχής αιτημάτων χρηστών μίγματος 2 για εικονικά δίκτυα πάνω σε μικρές, μεσαίες και μεγάλες τοπολογίες



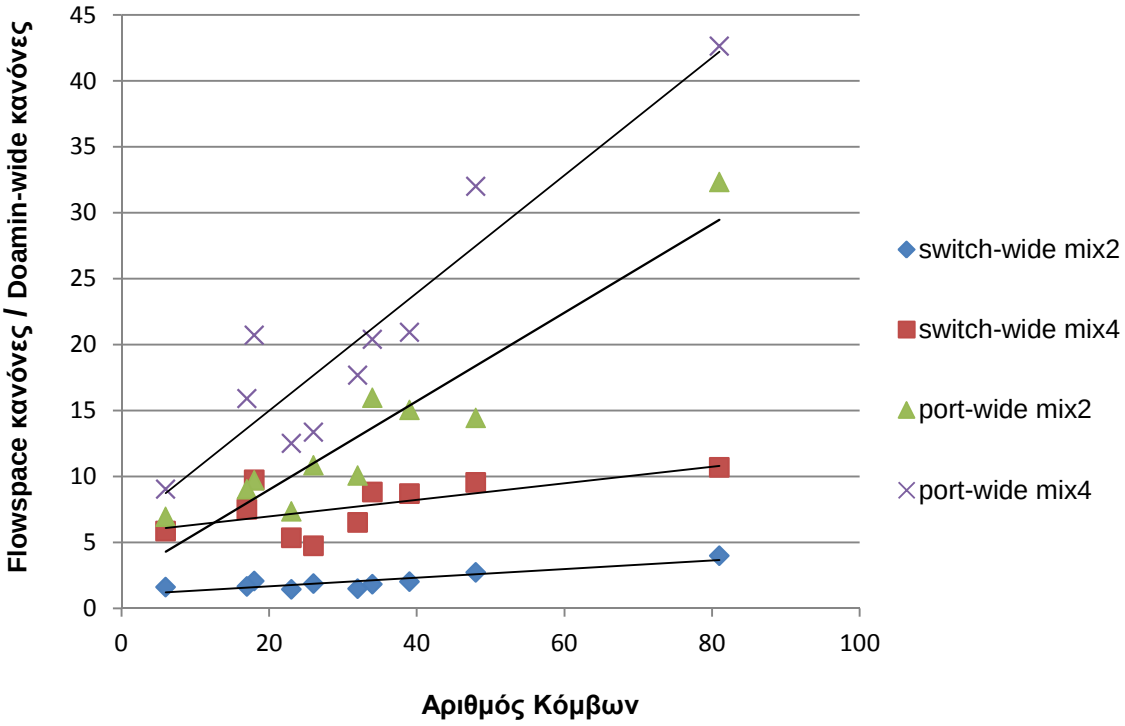
Σχήμα 29: Λόγος αποδοχής αιτημάτων χρηστών μίγματος 3 για εικονικά δίκτυα πάνω σε μικρές, μεσαίες και μεγάλες τοπολογίες

Στο **Σχήμα 28** και στο **Σχήμα 29** παρουσιάζουμε τους λόγους αποδοχής αιτημάτων χρηστών για το μίγμα 2 και το μίγμα 3 αντίστοιχα. Τα αποτελέσματα αφορούν το μέσο όρο των λόγων αποδοχής αιτημάτων για εικονικά δίκτυα πάνω σε τοπολογίες υποδομής που έχουν κατηγοριοποιηθεί βάσει του μεγέθους τους (μικρό, μεσαίο, μεγάλος), όπως αυτό έχει ορισθεί στην αρχή της Παραγράφου 6.5.2 . Ο αριθμός των αιτήσεων των χρηστών που έλαβε η FSP μηχανή για την εξαγωγή των αποτελεσμάτων στις διάφορες κατηγορίες τοπολογιών υποδομής κυμαίνεται (1.000, 4.000, 8.000, 16.000).

Μπορούμε να παρατηρήσουμε ότι δοθέντος μίγματος τοπολογιών για τα εικονικά δίκτυα που ζητούν οι χρήστες (**Σχήμα 28**, **Σχήμα 29)**, το χάσμα επίδοσης των port-wide και domain-wide μεθόδων μεγαλώνει όσο περισσότερα είναι τα αιτήματα. Επίσης, ανεξαρτήτως μίγματος, η domain-wide μέθοδος έχει σταθερή επίδοση είτε πρόκειται για μικρές τοπολογίες είτε για μεγάλες για συγκεκριμένο αριθμό αιτήσεων, αφού η δέσμευση του Packet ID γίνεται σε επίπεδο υποδομής, χωρίς να επηρεάζει το μέγεθος της υποδομής.

Παρατηρούμε ακόμα ότι για μεγάλες τοπολογίες και ειδικά για το μίγμα 3 η απόδοση της port-wide μεθόδου μειώνεται αργά, όσο αυξάνει ο αριθμός των συνολικών αιτήσεων. Ο λόγος είναι ότι σε μεγάλες τοπολογίες, όταν γίνεται η χρήση της μεθόδου port-wide slicing, υπάρχει μεγάλη πιθανότητα επαναχρησιμοποίησης του ίδιου Packet ID σε διαφορετικά σημεία της τοπολογίας υποδομής για την δέσμευση πόρων που θα αποδοθούν σε διαφορετικά εικονικά δίκτυα. Για αυτό το λόγο η port-wide μέθοδος είναι εξαιρετικά αποδοτική σε περιπτώσεις μεγάλων τοπολογιών που δέχονται και μεγάλο αριθμό αιτήσεων, ειδικά όταν αφήνεται στο επίπεδο διαχείρισης της υποδομής να αποφασιστεί ποιο Packet ID θα αποδοθεί ανά εικονικό δίκτυο.

Από τα αποτελέσματα που παρουσιάσαμε παραπάνω εξάγεται το συμπέρασμα ότι η switch-wide και ειδικά η port-wide μέθοδος έχουν πολύ καλή επίδοση συγκριτικά με τη domain-wide μέθοδο όσον αφορά το ποσοστό αποδοχής αιτήσεων για εικονικά δίκτυα. Ωστόσο, η απόδοσή τους έχει ως αντίτιμο τον μεγαλύτερο αριθμό απαιτούμενων κανόνων για τον ορισμό του flowspace.



Σχήμα 30: Αριθμός κανόνων για switch-wide και port-wide μεθόδους, κανονικοποιημένος ως προς domain-wide για 4.000 αιτήματα (mix2, mix4)

Στο **Σχήμα 30** παρουσιάζουμε τον αριθμό των απαιτούμενων κανόνων για τη switch-wide και port-wide μέθοδο κανονικοποιημένο προς τον αριθμό των απαιτούμενων κανόνων για τη domain-wide μέθοδο. Οι συγκεκριμένες μετρήσεις αφορούν 4.000 αιτήσεις για εικονικά δίκτυα, που πραγματοποιήθηκαν σε τοπολογίες υποδομής αποτελούμενες από 6 έως 81 κόμβους. Ο αριθμός των κόμβων των τοπολογιών υποδομής είναι η ανεξάρτητη μεταβλητή στο **Σχήμα 30**. Η παραγωγή των κανόνων έγινε βάσει των πολιτικών που περιγράφονται στην Παράγραφο 6.4.1 και της μεθόδου μείωσης κανόνων για τη switch-wide μέθοδο της Παραγράφου 6.4.2.

Τα αποτελέσματα αφορούν το μίγμα 2 και μίγμα 4 των αιτήσεων. Το μίγμα 2 βρέθηκε να είναι το χειρότερο σενάριο όσον αφορά την αύξηση των κανόνων, μεταξύ των 2 σεναρίων που αφορούσαν αποκλειστικά bound requests (mix1, mix2). Το μίγμα 4 είναι το σενάριο που παράγει τους περισσότερους κανόνες από κάθε άλλο (mix1, mix2, mix3, mix4), καθώς βάσει της πολιτικής που εφαρμόζεται με τα Packet IDs, είναι και το σενάριο που παρουσιάζει την μεγαλύτερη αποδοχή αιτημάτων για εικονικά δίκτυα.

Η αυξημένη αποδοχή αιτημάτων μεταφράζεται τελικά σε μεγάλο αριθμό κανόνων flowspace, σε σχέση με τον αριθμό κανόνων που απαιτούνται στην domain-wide μέθοδο που τα ποσοστά αποδοχής αιτημάτων είναι μικρά.

Επιπλέον, στο **Σχήμα 30** έχουμε τυπώσει τις ευθείες γραμμικής παλινδρόμησης για τις δυο (2) μεθόδους ανά μίγμα αιτήσεων. Μπορούμε να παρατηρήσουμε ότι η κλίση της ευθείας που αφορά την port-wide μέθοδο είναι μεγαλύτερη από αυτή που αφορά την switch-wide μέθοδο, με εντονότερο το φαινόμενο όταν έγινε χρήση του μίγματος 4. Το συγκεκριμένο αποτέλεσμα μπορεί να ερμηνευτεί αν αναλογιστούμε ότι στην domain-wide μέθοδο μια αίτηση χρήστη για εικονικό δίκτυο μεταφράζεται σε ένα κανόνα flowspace, ενώ στην switch-wide σε περισσότερους. Τέλος, η port-wide μέθοδος βάσει της ανάλυσης που έγινε στην Παράγραφο 6.4.2., οδηγεί στη δημιουργία ακόμα περισσότερων κανόνων.

Το τελευταίο στάδιο της αξιολόγησης περιελάμβανε μετρήσεις για τη χρονική επιβάρυνση που εισάγεται στο επίπεδο ελέγχου μεταξύ των OpenFlow switches και των OpenFlow controllers των χρηστών, λόγω του επιπέδου διαμοιρασμού ελέγχου (slicing layer of the control-plane). Συγκεκριμένα, όλα τα μηνύματα ελέγχου από και προς τα OpenFlow switches περνούν από το συγκεκριμένο επίπεδο για να προωθηθούν στον αρμόδιο OpenFlow controller του χρήστη, βάσει των κανόνων διαμοιρασμού του flowspace.

Οι μετρήσεις έγιναν με χρήση του FlowVisor (έκδοση 1.4), ο οποίος παρουσιάζει μεγάλη βελτίωση στους χρόνους αναζήτησης των κανόνων διαμοιρασμού του flowspace. Η συγκεκριμένη έκδοση βελτιώνει κατά πολύ τους χρόνους αναζήτησης καθώς υλοποιεί αλγορίθμους κατατεμαχισμού (hashing) που περιγράφονται στο [Knuth97], αντί για απλή γραμμική αναζήτηση (linear search) που είχε στις πρώτες υλοποιήσεις [Sher10] .

Οι flowspace κανόνες βάσει της port-wide μεθόδου διαμοιρασμού που δημιουργήθηκαν στην μηχανή FSP και εισήχθησαν στον FlowVisor αφορούσαν τυχαία αιτήματα του μίγματος 2 για εικονικά δίκτυα πάνω στην τοπολογία υποδομής του GÉANT (**Πίνακας 6-3)**. Χρησιμοποιήθηκε η συγκεκριμένη μέθοδος διαμοιρασμού, δεδομένου ότι είναι η πιο απαιτητική από άποψη αριθμού κανόνων, όπως έχουμε αναλύσει παραπάνω.

Για τη μέτρηση της χρονικής επιβάρυνσης που εισάγει ο proxy controller, σημειώναμε τον χρόνο εισόδου των πακέτων ελέγχου που προέρχονταν από τα OpenFlow switches στο southbound interface του και τον χρόνο εξόδου από το northbound interface του, με κατεύθυνση τους controllers των χρηστών, χρησιμοποιώντας την βιβλιοθήκη [Libpcap]. Η διαφορά των δυο χρόνων ήταν ίση με την συνολική επιβάρυνση που εισήγαγε ο proxy controller ως σύστημα. Η συγκεκριμένη επιβάρυνση περιλαμβάνει, εκτός από το χρόνο αναζήτησης των flowtables που περιέχουν τους κανόνες του flowspace, όλες εκείνες τις διεργασίες που απαιτούνται από το λογισμικό του proxy controller καθώς και από το επίπεδο του λειτουργικού συστήματος για τον χειρισμό των OpenFlow λειτουργιών.

Παρουσιάζονται τα αποτελέσματα για τον αριθμό κανόνων, τη χρονική επιβάρυνση καθώς και η κατανάλωση μνήμης στον εξυπηρετητή που υλοποιεί τις λειτουργίες του proxy controller για την δημιουργία, την ανανέωση, την διαγραφή και τον χειρισμό των κανόνων του flowspace (**Πίνακας 6-3).** Όπως παρατηρούμε από τα αποτελέσματα, ο FlowVisor εισάγει μικρή χρονική καθυστέρηση η οποία μάλιστα λόγω των hashing αλγορίθμων αναζήτησης δεν αυξάνει σημαντικά με την αύξηση των κανόνων του flowspace, ακόμα και στην περίπτωση που ο συνολικός αριθμός κανόνων αγγίζει τις 180.000. Σε περιπτώσεις που οι αιτήσεις χρηστών είναι 7.000 παρατηρούμε ότι η κατανάλωση μνήμης φτάνει τα 7,5 Gbyte. Η συγκεκριμένη τιμή, αν και αρκετά μεγάλη, δεν είναι απαγορευτική για ένα σύγχρονο εξυπηρετητή.

Ο χρόνος εισαγωγής νέων flowspace κανόνων στον FlowVisor από την FSP μηχανή ήταν σταθερός και δεν αύξανε όταν οι ήδη εγκαθιδρυμένοι κανόνες είχαν μεγάλο πλήθος. *Συνεπώς, το μεγάλο πλήθος εγκαθιδρυμένων κανόνων για το flowspace μέσα στο FlowVisor δεν δημιουργούσε πρόβλημα εισαγωγής κανόνων για νέα εικονικά δίκτυα που το επίπεδο διαχείρισης δικτύου ήθελε να δημιουργήσει για την εξυπηρέτηση νέων πελατών.*

	Τοπολογία υποδομής GÉANT		
Αιτήσεις χρηστών (mix2)	**Αριθμός κανόνων (port-wide μέθοδος)**	**Χρονική επιβάρυνση (ms)**	**Κατανάλωση μνήμης (Mbytes)**
1.000	17K	4.7	672
2.000	45K	5.0	1852
4.000	100K	5.4	4050
6.000	150K	5.5	6102
7.000	180K	5.9	7500

Πίνακας 6-3: Χρονική επιβάρυνση/κατανάλωση μνήμης από το επίπεδο διαμοιρασμού του flowspace (υλοποιημένο με FlowVisor) στην τοπολογία του GÉANT

7 Συμπεράσματα και Μελλοντική Έρευνα

7.1 Συμπεράσματα

Τα αποτελέσματα της παρούσας Διδακτορικής Διατριβής συνοψίζονται με κριτήριο το επίπεδο λειτουργίας δικτύου στο οποίο συνέβαλε:

- **Επίπεδο προώθησης πακέτων (data-plane)**

Στην παρούσα εργασία έγινε δυνατή η δικτυακή διασύνδεση εικονικών υπολογιστικών οντοτήτων (virtual node) ετερογενών ομόσπονδων υποδομών με σκοπό την ταυτόχρονη εξυπηρέτηση πολλαπλών χρηστών. Η λύση βασίστηκε στη δημιουργία λογικών διεπαφών GRE και IEEE 802.1Q στο επίπεδο προώθησης πόρων, με το επίπεδο διαχείρισης δικτύου να φροντίζει για την απομόνωση των εικονικών δικτύων. Ο μηχανισμός που αναπτύχθηκε μπορεί να επιτύχει διασύνδεση εικονικών υπολογιστικών υποδομών Layer 2 και Layer 3 επιτυγχάνοντας κλιμακοθετησιμότητα, ξεπερνώντας τους περιορισμούς που εισάγει η χρήση VLANs σε επιμέρους τοπολογίες. Δίνει επίσης την ευκαιρία ευέλικτης δημιουργίας/καταστροφής εικονικών τοπολογιών από το ομοσπονδοποιημένο επίπεδο διαχείρισης δικτύου, χωρίς να εισάγεται ιδιαίτερη χρονική επιβάρυνση στην επικοινωνία των κόμβων ή μείωση της ταχύτητας των εικονικών ζεύξεων.

- **Επίπεδο διαχείρισης (management-plane)**

Έγινε μελέτη των εγγενών δυνατοτήτων του OpenFlow για παρακολούθηση και δημιουργία μηχανισμού που συνδυάζει τα πρωτόκολλα OpenFlow και sFlow για αποδοτικότερη συλλογή στοιχείων παθητικής παρακολούθησης σε περιβάλλοντα SDN.

Οι πειραματικές μετρήσεις που έγιναν σε πρωτότυπο μηχανισμού αντιμετώπισης δικτυακών ανωμαλιών δείχνει ότι η συλλειτουργία του sFlow και OpenFlow είναι αποτελεσματική ακόμα και για απαιτητικές εφαρμογές πραγματικού χρόνου, όπως είναι ο εντοπισμός και η εξομάλυνση της δικτυακής κίνησης υπό συνθήκες δικτυακών επιθέσεων. Ο προτεινόμενος μηχανισμός προσφέρει μείωση της

επικοινωνίας στο κεντρικοποιημένο επίπεδο ελέγχου των οριζομένων από λογισμικό δικτύων (δίκτυα OpenFlow), μείωση της απαιτούμενης επεξεργαστικής ισχύος σε OpenFlow switches/controllers, καθώς και μείωση (μια τάξη μεγέθους) του αριθμού των εγγραφών ροών μέσα στα OpenFlow switches. Η σημαντικότερη συνδρομή του προτεινόμενου μηχανισμού είναι η αποφυγή υπερφόρτωσης του επιπέδου ελέγχου από ανωμαλίες που μπορεί να προκύψουν στο επίπεδο προώθησης δεδομένων.

Για τους παραπάνω λόγους, προτείνουμε το πλαίσιο παρακολούθησης «Passive Flow Monitoring» (PaFloMon), για την κάλυψη των αναγκών παθητικής παρακολούθησης της δικτυακής κίνησης. Το PaFloMon εξασφαλίζει ορθή λειτουργία των κεντρικοποιημένων λειτουργιών δικτύωσης λόγω της ταυτόχρονης χρήσης του OpenFlow με πρωτόκολλα όπως το sFlow και το SNMP. Το PaFloMon έχει κλιμακοθετησιμότητα σε μεγάλο εύρος hardware και software μεταγωγέων (π.χ Open vSwitch) που χρησιμοποιούνται κατά κόρον σε περιβάλλοντα εικονικοποίησης.

Για περιβάλλοντα πολλαπλών χρηστών/ενοικιαστών προτείνεται η δυνατότητα αποθήκευσης, κατηγοριοποίησης και εμφάνισης των στατιστικών στοιχείων ανά slice για λόγους ιδιωτικότητας (privacy) και απομόνωσης (isolation) των χρηστών/ενοικιαστών της υποδομής. Οι πειραματικές μετρήσεις έδειξαν ότι σε ένα διαμοιραζόμενο εικονικό περιβάλλον με OpenFlow switches επιβάλλεται ο περιορισμός ανά χρήστη του αριθμού των flow entries που μπορεί να εισάγει σε κάθε OpenFlow switch. Στη περίπτωση του PaFloMon δίνεται η δυνατότητα να επιβληθεί μια πολιτική όσον αφορά τον αριθμό των OpenFlow entries ανά switch ανά χρήστη μέσω μιας εκ των δυο αρχιτεκτονικών διαμοιρασμού πόρων του επιπέδου ελέγχου που περιγράφονται παρακάτω.

- **Επίπεδο ελέγχου (control-plane)**

Διαμορφώθηκαν δυο αρχιτεκτονικές για τη δημιουργία πολλαπλών εικονικών δικτύων οριζόμενα από λογισμικό που διαμοιράζονται την ίδια δικτυακή υποδομή.

Η πρώτη αρχιτεκτονική επιτρέπει την ανάθεση μοναδικής περιοχής του συνολικού πεδίου ορισμού ροών του OpenFlow (flowspace), όπως αυτό ορίζεται από το OpenFlow πρωτόκολλο, σε κάθε εικονικό δίκτυο χρήστη. Ο OpenFlow controller του χρήστη ελέγχει συγκεκριμένο μέρος flowspace πάνω σε ένα τμήμα

της τοπολογίας υποδομής που του εκθέτει ένας ενδιάμεσος OpenFlow proxy controller (υπεύθυνος για το διαμοιρασμό του flowspace).

Η δεύτερη αρχιτεκτονική επιτρέπει την έκθεση ενός εικονικού δικτύου σε OpenFlow controllers χρηστών. Το αποτελούμενο από εικονικούς πόρους (κόμβους, ζεύξεις) δίκτυο αποκρύπτει την πραγματική τοπολογία υποδομής και δίνει την ικανότητα για λειτουργίες τύπου path splitting και path reservation.

Οι δυο αρχιτεκτονικές μπορούν να διαμοιράζουν το διαθέσιμο flowspace βάσει μιας πολιτικής που έχει επιλέξει ο διαχειριστής υποδομής, με στόχο να διατηρεί την απομόνωση μεταξύ των εικονικών δικτύων.

Προτείνουμε και μελετούμε τρεις μεθόδους διαμοιρασμού του flowspace για την εφαρμογή πολιτικής διαμοιρασμού του flowspace με γνώμονα την απόλυτη απομόνωση μεταξύ των εικονικών δικτύων των χρηστών στο επίπεδο ελέγχου αλλά και στο επίπεδο προώθησης δεδομένων: (i) **domain-wide**, (ii) **switch-wide**, (iii) **port-wide**. Συμπεραίνουμε ότι η domain-wide πολιτική που κάνει των διαμοιρασμό του flowspace σε επίπεδο υποδομής απαιτεί ένα μικρό αριθμό κανόνων εφαρμογής της πολιτικής, με αντίτιμο τον μικρό λόγο αποδοχής αιτημάτων χρηστών για εγκαθίδρυση νέων εικονικών δικτύων. Αντίθετα, η port-wide μέθοδος που στον διαμοιρασμό του flowspace λαμβάνει υπόψη και τις δικτυακές πόρτες που θέλει να χρησιμοποιήσει κάθε εικονικό δίκτυο παρουσιάζει τα καλύτερα αποτελέσματα όσον αφορά την ικανοποίηση των αιτήσεων των χρηστών, με αντίτιμο τους περισσότερους κανόνες υλοποίησης της πολιτική διαμοιρασμού. Η switch-wide μέθοδος μπορεί να αποτελέσει μια ενδιάμεση λύση ως προς την ικανοποίηση αιτημάτων για νέα εικονικά δίκτυα. Για τη switch-wide μέθοδο προτείνουμε χρήση κανόνων ειδικού σκοπού (aggregated rules) και ενός OpenFlow controller υποδομής για μείωση των κανόνων επιβολής της πολιτικής διαμοιρασμού και απομόνωσης.

Επιπλέον εισάγουμε την έννοια της μηχανής διαμοιρασμού του flowspace (**Flowspace slicing policy - FSP**) που δύναται να εφαρμόζει μια από τις τρεις μεθόδους. Η συγκεκριμένη μηχανή είναι υπεύθυνη για να μετατρέπει τις πολιτικές του επιπέδου διαχείρισης της υποδομής (substrate management-plane policies) σε κανόνες ελέγχου ροών των εικονικών δικτύων και υλοποιήθηκε ως πρωτότυπο. Η χρήση του πρωτοτύπου δοκιμάσθηκε σε συνδυασμό με proxy controller (FlowVisor) που εφήρμοζε τους κανόνες που παράγονταν για WAN τοπολογίες

όπως το Internet2/OS3E και το GÉANT. Οι πειραματικές μετρήσεις δείχνουν ότι ακόμα και η port-wide μέθοδος που απαιτεί μεγαλύτερο αριθμό πόρων στο επίπεδο διαμοιρασμού ελέγχου (π.χ. μνήμη συστήματος) δεν εισάγει μεγάλη χρονική επιβάρυνση στο κεντρικοποιημένο επίπεδο ελέγχου, ακόμα και όταν τα αιτήματα για εικονικά δίκτυα είναι πάνω από 15.000.

7.2 Μελλοντική Έρευνα

Οι προεκτάσεις της παρούσας Διδακτορικής Διατριβής και τα πιθανά ερευνητικά πεδία κατηγοριοποιούνται βάσει του επιπέδου λειτουργίας δικτύου:

- **Επίπεδο προώθησης πακέτων (data-plane)**

Στην παρούσα εργασία έγινε δυνατή η δικτυακή διασύνδεση εικονικών υπολογιστικών οντοτήτων (virtual node) ετερογενών ομόσπονδων υποδομών, με σκοπό την ταυτόχρονη εξυπηρέτηση πολλαπλών χρηστών. Η λύση βασίστηκε στη δημιουργία λογικών διεπαφών στο επίπεδο προώθησης πόρων, με το επίπεδο διαχείρισης δικτύου να φροντίζει για την απομόνωση των εικονικών δικτύων. Η συγκεκριμένη τακτική επιτυγχάνει τη διασύνδεση, αλλά δεν δίνει τη δυνατότητα στους χρήστες να επιλέγουν μηχανισμό προώθησης βασισμένο στην λογική προώθησης ροών που έχει εισάγει το OpenFlow. Ο μηχανισμός των Layer 2 Ethernet switches λογισμικού (Linux bridges, OVS χωρίς χρήση OpenFlow) υλοποιεί μόνο τη MAC learning and forwarding λογική προώθησης. Επέκταση της εργασίας αποτελεί η δημιουργία μηχανισμών προώθησης ροών ελεγχόμενων από το χρήστη, με τη βοήθεια του πρωτοκόλλου OpenFlow, σε ένα περιβάλλον ομόσπονδων περιοχών.

- **Επίπεδο διαχείρισης (management-plane)**

Το πλαίσιο παθητικής παρακολούθησης υποδομών δικτυακής εικονικοποίησης PaFloMon, που παρουσιάζουμε στην εργασία μας, καλύπτει τις ανάγκες παρακολούθησης δικτυακής κίνησης προσφέροντας ατροσία του επιπέδου ελέγχου από δικτυακές ανωμαλίες του επιπέδου προώθησης δεδομένων. Η ανάλυση των

πολιτικών παθητικής παρακολούθησης και του τρόπου υλοποίησής τους σε ένα περιβάλλον πολλαπλών χρηστών/ενοικιαστών, πολλαπλών εικονικών δικτύων δεν αποτέλεσε μέρος της εργασίας μας, εντούτοις παραμένει ανοιχτό θέμα σε εικονικές υποδομές πολλαπλών χρηστών.

Οι μηχανισμοί που κάνουν χρήση του OpenFlow και sFlow μπορεί να έχουν ευεργετικά αποτελέσματα όσον αφορά την κλιμακοθετησιμότητα απαιτητικών λειτουργιών, όπως αυτή του εντοπισμού δικτυακών ανωμαλιών, χωρίς όμως να έχει μελετηθεί πως μπορεί να γίνει ανταλλαγή πληροφοριών και εξομάλυνση δικτυακών ανωμαλιών μεγάλης κλίμακας σε ομόσπονδες υποδομές δικτύων που ελέγχονται από λογισμικό (federated SDNs), σημείο άξιο της ερευνητικής προσοχής.

- **Επίπεδο ελέγχου (control-plane)**

Οι μέθοδοι διαμοιρασμού του flowspace, που προτάθηκαν και εφαρμόσθηκαν στο πρωτότυπο της μηχανής διαμοιρασμού του flowspace (Flowspace slicing policy - FSP), αφορούν μια διαχειριστική περιοχή, ελεγχόμενη από ένα ενιαίο επίπεδο διαχείρισης υποδομής. Η συγκεκριμένη εργασία δύναται να επεκταθεί με μελέτη και ανάπτυξη μηχανισμών που υποστηρίζουν την εφαρμογή πολιτικών διαμοιρασμού flowspace σε ένα διατομεακό διαχειριστικό περιβάλλον. Η συγκεκριμένη προέκταση της Διδακτορικής Διατριβής θα απαιτούσε ανταλλαγή πληροφοριών διαμοιρασμού του επιπέδου ελέγχου μεταξύ διαφορετικών διαχειριστικών περιοχών, λαμβάνοντας υπόψη τις απαιτήσεις των παρόχων υποδομής για απόκρυψη ευαίσθητων πληροφοριών καθώς και μηχανισμούς συνάθροισης και αφαίρεσης για την επιτυχημένη κλιμακοθετησιμότητα των λύσεων.

8 Βιβλιογραφία

[Ahme07] T. Ahmed, B. Oreshkin, M. Coates, Machine learning approaches to network anomaly detection, in SYSML'07 Proceedings of the 2nd USENIX workshop on Tackling computer systems problems with machine learning techniques , 2007

[Andr09] G. Androulidakis, V. Chatzigiannakis and S. Papavassiliou, “Network Anomaly Detection and Classification via Opportunistic Sampling”, IEEE Network, Vol. 23, No. 1, pp. 6-12, 2009

[Argy13] Argyropoulos, C.; Androulidakis, G.; Kalogeras, D.; Pietrzak, B.; Belter, B.; Lymberopoulos, L.; Maglaris, V., "Network virtualization over heterogeneous federated infrastructures: Data plane connectivity,"*Integrated Network Management (IM 2013), 2013 IFIP/IEEE International Symposium on* , vol., no., pp.26,33, 27-31, 2013

[Bavie06] Bavier, A., Feamster, N., Huang, M. , Peterson, L. , Rexford J., “In VINI veritas: realistic and controlled network experimentation”, in Proc. of the 2006 conference on Applications, technologies, architectures, and protocols for computer communications, September, Pisa, Italy, 2006

[Bhan97] R. Bhandari, Optimal physical diversity algorithms and survivable networks, Proc. Second IEEE Symposium on Computers and Communications 1997, Alexandria, Egypt, pp. 433-441, 1997

[Bhat08] Bhatia, S., Motiwala, M., Muhlbauer, W., Mundada, Y., Valancius, V., Bavier, A., Feamster, N., Peterson, L., Rexford, J., “Trellis: a platform for building flexible, fast virtual networks on commodity hardware”, in Proceedings of the ACM CoNEXT, 2008, Conference (CoNEXT '08),

ACM, New York, NY, USA, Article 72

[Blan12] E. Blanton, S. C Chatterjee, S. Gangam, S. Kala, D. Sharma, S. Fahmy, P. Sharma, “Design and Evaluation of the S3 Monitor Network Measurement Service on GENI, Fourth International Conference on Communication Systems and Networks (COMSNETS)”, 2012

[Brag10] R. Braga, E. Mota, A. Passito, Lightweight DDoS flooding attack detection using NOX/OpenFlow, in LCN '10 Proceedings of the 2010 IEEE 35th Conference on Local Computer , 2010, pp. 408-415

[Casa07] M. Casado, M. J. Freedman, J. Pettit, J. Luo, N. McKeown, and S. Shenker, “Ethane: Taking Control of the Enterprise”, In Proc. SIGCOMM, August 2007

[Casa10] Martin Casado, Teemu Koponen, Rajiv Ramanathan, and Scott Shenker, 2010, “Virtualizing the network forwarding plane”, In Proceedings of the Workshop on Programmable Routers for Extensible Services of Tomorrow (PRESTO '10), ACM, New York, NY, USA

[Chow12] M. Chowdhury, M. R. Rahman, R. Boutaba, ViNEYard: virtual network embedding algorithms with coordinated node and link mapping, IEEE/ACM Transactions on Networking, vol. 20, Issue 1, pp. 206-219, 2012

[Chun03] Chun, B., Culler, D., Roscoe, Bavier, A., Peterson, L., Wawrzoniak, M., Bowman M., “PlanetLab: an overlay testbed for broad-coverage services”, SIGCOMM Comput. Commun. Rev. 33, 3, 2003

[Cisco] Cisco White Paper, “The Benefits of Centralization in Wireles LANs via the Cisco Unified Wireless Network”

[Deri00] L. Deri, S. Suin, Effective traffic measurement using ntop, *IEEE Communications Magazine*, vol. 38, no. 5, pp. 138-143, May 2000.

[DMTF] DMTF, Distributed Management Task Force, www.dmtf.org

[Dori12] R. Doriguzzi Corin, M. Gerola, R. Riggio, F. DePellegrini, E. Salvadori, VeRTIGO: Network Virtualization and Beyond, Software Defined Networking (EWSDN), 2012 European Workshop on SDN, vol., no., pp.24, 29, 2012

[Drut13] D. Drutskoy, E. Keller, J. Rexford, Scalable Network Virtualization in Software-Defined Networks, Internet Computing, IEEE, vol.17, no.2, pp.20,27, 2013

[Emulab] Emulab, http://www.emulab.net

[FEDERICA] FEDERICA – Federated E-infrastructure Dedicated to European Researchers Innovating in Computing network Architectures, http://www.fp7-federica.eu

[FIRE] http://cordis.europa.eu/fp7/ict/fire/

[FlowFire] Flowspace Firewall, http://globalnoc.iu.edu/software/sdn.html

[FOAM] https://openflow.stanford.edu/display/FOAM/Home

[FSPengine] Flowspace Slicing Policy engine, https://github.com/spirosmastorakis/FSP_Engine

[GEANT] GÉANT, http://www.geant.net

[GEANTOF] Technical Annex B –GÉANT OpenFlow Facility, http://www.geant.net

[GEANTtopo] GÉANT, the pan-European research and education network that interconnects Europe's National Research and Education Networks (NRENs).

http://www.geant.net/Resources/Media_Library/Pages/Maps.aspx

[GEMINI] "GEMINI Instrumentation and Measurement Tool", http://groups.geni.net/geni/wiki/GEMINITutorial

[GENI] http://www.geni.net/

[GENIIM] "The Global Environment for Network Innovations (GENI) Instrumentation and Measurement Work in Progress (I&M)", http://groups.geni.net/geni/wiki/GeniInstMeas

[GENIIN] "Instrumentation tools for a GENI prototype", http://groups.geni.net/geni/wiki/InstrumentationTools

[GENION] "OnTimeMeasure: Centralized and distributed measurement orchestration software", http://groups.geni.net/geni/wiki/OnTimeMeasure

[Giot14] K. Giotis, C. Argyropoulos, G. Androulidakis, D. Kalogeras, V. Maglaris, Combining OpenFlow and sFlow for an effective and scalable anomaly detection and mitigation mechanism on SDN environments, Computer Networks, Volume 62, 7 April 2014, Pages 122-136, ISSN 1389-1286

[GRNOC] Global Research Network Operations Center, http://globalnoc.iu.edu

[Gupt01] Gupta, P.; McKeown, N., "Algorithms for packet classification," *Network, IEEE* , vol.15, no.2, pp.24,32, Mar/Apr 2001

[Hami09] J. Hamilton, "Data center networks are in my way", Talk at Stanford Clean Slate CTO Summit, 2009

[Hand05] Handley, M., Kohler, E., Ghosh, A., Hodson, O., Radoslavov, P., "Designing extensible IP router software" , Proc. Networked Systems Design and Implementation , 2005

[Hell12] B. Heller, R. Sherwood, N. McKeown, The controller placement problem, in Proceedings of the first workshop on Hot topics in Software Defined Networks (HotSDN '12). ACM, New York, USA, pages 7-12, 2012

[Hible08] Hibler, M., Ricci, R., Stoller, L., Duerig, J., Guruprasad, S., Stack, T., Webb, K., Lepreau, J., “Large-scale virtualization in the Emulab network testbed. In USENIX 2008 Annual Technical Conference on Annual Technical Conference (ATC'08)”, USENIX Association, Berkeley, CA, USA, 113-128

[Hjal00] Hjalmtysson, G., "The Pronto platform: a flexible toolkit for programming networks using a commodity operating system," *Open Architectures and Network Programming, Proceedings OPENARCH 2000, IEEE Third Conference on* , vol., no., pp.98,107, Mar 2000

[Huan05] Huang, M., “VNET: PlanetLab Virtualized Network Access” , Tech. Rep. PDN–05–029, PlanetLab Consortium, 2005

[IN2topo] Internet2 Open Science, Scholarship and Services exchange http://inddi.wikispaces.com/Internet2-based+NDDI

[INtopo] “The Internet Topology Zoo” project, http://www.topology-zoo.org

[Ioan00] S. Ioannidis, A. D. Keromytis, S. M. Bellovin, and J. M.Smith, Implementing a Distributed Firewall. In Proc. CCS, 2000

[ITU-I322] ITU-T, I.322, SERIES I: INTEGRATED SERVICES DIGITAL NETWORK, Overall network aspects and functions – Refe rence models, Generic protocol reference model for telecommunication networks, http://www.itu.int/rec/T-REC-I.322-199902-I/en

[JUNOS] JUNOS Logical Systems,

http://www.juniper.net/techpubs/en_US/junos/information-products/pathway-pages/config-guide-logical-systems/config-guide-logical-systems.html#overview

[Knuth97] D. Knuth, The Art of Computer Programming, Volume 3: Sorting and Searching, Second Edition, Chapter 6.3, page 492. Addison Wesley, 1997

[Kohl00] Kohler, E., Morris, R., Chen, B., Jannotti, J., Kaashoek M. F., “The Click modular router”, ACM Transactions on Computer Systems, vol. 18, pp. 263–297, 2000

[Kopo10] T. Koponen, M. Casado, N. Gude, J. Stribling, L. Poutievski, M. Zhu, R. Ramanathan, T. Hama, and S. Shenker, Onix: A Distributed Control Platform for Large-scale Production Networks, in Proceedings of USENIX OSDI, page 351-364, 2010

[Kout04] Koutsonikola, V.; Vakali, A., "LDAP: framework, practices, and trends," Internet Computing, IEEE , vol.8, no.5, pp.66,72, Sept.-Oct. 2004

[Lakh05] A. Lakhina, M. Crovella, and C. Diot, Mining anomalies using traffic feature distributions, in ACM SIGCOMM, 2005, pp. 217 - 228

[LAMP] “Leveraging and abstracting measurements with perfSONAR (LAMP)”, http://groups.geni.net/geni/wiki/LAMP

[Libpcap] Tcpdump/Libpcap, http://www.tcpdump.org

[LXC] LXC - The userspace control package for Linux Containers, http://lxc.sourceforge.net

[Lymp12] L. Lymberopoulos, M. Grammatikou, M. Potts, P. Grosso, A. Fekete, B. Belter, M. Campanella and V. Maglaris, "NOVI Tools and Algorithms for Federating Virtualized Infrastructures", Future Internet – From Technological Promises to Reality, Springer Lecture Notes in Computer Science, pp. 213-224, 2012

[M3400] "M.3400 TMN management functions", International Telecommunications Union, 1997

[McGe65] McGee, W. C., "On dynamic program relocation," IBM Systems Journal , vol.4, no.3, pp.184,199, 1965

[McKe08] Nick McKeown, Tom Anderson, Hari Balakrishnan, Guru Parulkar, Larry Peterson, Jennifer Rexford, Scott Shenker, and Jonathan Turner, 2008, OpenFlow: enabling innovation in campus networks. SIGCOMM Comput. Commun. Rev. 38, 2 (March 2008), 69-74

[Mehd11] S. Mehdi, J. Khalid, S. Khayam, Revisiting traffic anomaly detection using software defined networking, in *RAID'11 Proceedings of the 14th international conference on Recent Advances in Intrusion Detection*, 2011, pp. 161-180

[NetFlow] Ed. B. Claise, "Cisco Systems NetFlow Services Export Version 9," *RFC 3954*, October 2004

[NEXUS] Cisco Nexus 1000V Switch, http://www.cisco.com/en/US/products/ps9902/index.html

[NMKA] A. Farrel et al, Network Management Know It All, Morgan Kaufmann, ISBN-13: 978-0123745989, 2011

[NOVI] NOVI – Network Innovation over Virtualized Infrastructures FP7 STREP Project, http://www.fp7-novi.eu

[NOX] N. Gude, T. Koponen, J. Pettit, B. Pfaff, M. Casado, N. McKeown, S. Shenker, NOX: towards an operating system for networks, SIGCOMM Computer Communication Review 38, 3, 105-110, 2008

[NTT] NTT Communications' Enterprise Cloud Goes Global

http://www.ntt.com/aboutus_e/news/data/20130220.html

[OCF] OFELIA Control Framework, http://www.fp7-ofelia.eu/assets/Public-Deliverables/OFELIAD41.pdf

[OFELIA] OFELIA, http://www.fp7-ofelia.eu

[OFMF] “OpenFlow Monitoring Facility”, https://openflow.stanford.edu/display/SDEP/Management

[OFspec10] OpenFlow Switch Specification 1.0.0 https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow/openflow-spec-v1.0.0.pdf

[OFspec14] OpenFlow Switch Specification 1.4.0, https://www.opennetworking.org/sdn-resources/onf-specifications/openflow

[OneLab] OneLab, http://www.onelab.eu

[ONF12] ONF White Paper, “Software-Defined Networking: The New Norm for Networks”, 2012

[OpenVPN] OpenVPN, https://www.openvpn.net

[OSI] ISO/IEC 7498-4: Information processing systems -- Open Systems Interconnection -- Basic Reference Model -- Part 4: Management framework

[OVS] Open vSwitch, http://openvswitch.org

[OVSDB] B. Pfaff, B. Davie, Ed., "The Open vSwitch Database Management Protocol", IETF draft-pfaff-ovsdb-proto-03, 2013

[PanLab] PanLab, http://www.panlab.net

[Papa13] C. Papagianni, A. Leivadeas, S. Papavassiliou, V. Maglaris, C. Cervello-Pastor, A. Monje, On the optimal allocation of virtual resources in cloud computing networks, IEEE Transactions on Computers, vol. 62, no. 6, pp. 1060–1071, 2013

[Patc07] A. Patcha, J.-M. Park, An overview of anomaly detection techniques: Existing solutions and latest technological trends, Computer Networks: The International Journal of Computer and Telecommunications Networking, vol. 51, no. 12, pp. 3448-3470 , August 2007

[PBNM] J. Strassner, Policy-based Network Management: Solutions for the Next Generation, Morgan Kaufmann, ISBN-13: 978-1558608597, 2003

[perfSONAR] perfSONAR, http://www.perfsonar.net/

[Perr05] H.G. Perros, Connection-Oriented Networks: SONET/SDH, ATM, MPLS and Optical Networks, John Wiley & Sons, 2005

[Pete00] L. Peterson, Y. Gottlieb, M. Hibler, P. Tullmann, J. Lepreau, S. Schwab, H. Dandekar, A. Purtell, and J. Hartman, “A NodeOS interface for active networks”, IEEE Journal of Selected Areas in Communications, March 2001

[Pfaf09] Pfaff, B., Pettit, J., Amidon, K., Casado, M., Koponen, T., Shenker, S., “Extending Networking into the Virtualization Layer”, HotNets, 2009

[PlanetLabH] PlanetLab history, http://www.planet-lab.org/history

[ProtoGeni] ProtoGeni, http://www.protogeni.net

[Rexf04] J. Rexford, A. Greenberg, G. Hjalmtysson, et al., In Proceedings of HotNets 2004 (November 2004) Network-Wide Decision Making: Toward A Wafer-Thin Control Plane

[RFC2784] Farinacci, D., Li, T., Hanks, S., Meyer, D., Traina P., “Generic routing encapsulation (GRE)”, IETF RFC 2784, 2000

[RFC2890] Dommety, G., “Key and Sequence Number Extensions to GRE”, IETF RFC 2890, 2000

[RFC3198] Terminology for Policy-Based Management, RFC 3198 (Informational), http://datatracker.ietf.org/doc/rfc3198

[RFC3416] R. Presuhn et al., Version 2 of the Protocol Operations for the Simple Network Management Protocol (SNMP), IETF RFC 3416, Internet standard, 2002

[RFC3460] Policy Core Information Model (PCIM) Extensions, RFC3460 RFC (Proposed Standard), http://datatracker.ietf.org/doc/rfc3460

[RFC3641] S. Legg , Generic String Encoding Rules (GSER) for ASN.1 Types, IETF RFC 3641, Proposed Standard, 2003

[RFC3746] Yang, L., Dantu, R., Anderson, T., and R. Gopal, "Forwarding and Control Element Separation (ForCES) Framework", RFC 3746, April 2004

[RFC3954] B. Claise, Ed., Cisco Systems NetFlow Services Export Version 9, IETF RFC 3954, Informational, 2004

[RFC4506] M. Eisler et al , “XDR: External Data Representation Standard”, IETF RFC 4506, Internet Standard, 2006

[RFC5951] Lam, K., Mansfield, S., and E. Gray "Network Management Requirements for MPLS-based Transport Networks" RFC 5951, September 2010

[Rigg13] R. Riggio, F. De Pellegrini, E. Salvadori, M. Gerola, C. R. Doriguzzi Corin. Progressive virtual topology embedding in OpenFlow networks, Integrated Network Management (IM 2013), IFIP/IEEE International Symposium on Integrated Network Management, Ghent, vol., no., pp.1122,1128, 2013

[RSpec] ProtoGENI RSpec, http://www.protogeni.net/trac/protogeni/wiki/RSpec

[Scapy] Scapy, http://www.secdev.org/projects/scapy

[sFlow] P. Phaal, sFlow Specification Version 5

[sFlow802.11] sFlow 802.11 structures, http://www.sflow.org/sflow_80211.txt

[sFlowHost] Host sFlow, http://host-sflow.sourceforge.net

[Shah07] S. Shah, A. Nucci, M. Munafo, R. Cruz, and S. Muthukrishnan, DoWitcher: Effective Worm Detection and Containment in the Internet Core, in *INFOCOM 2007. 26th IEEE International Conference on Computer Communications. IEEE*, May 2007, pp. 2541 – 2545

[Sher10] Rob Sherwood, Glen Gibb, Kok-Kiong Yap, Guido Appenzeller, Martin Casado, Nick McKeown, and Guru Parulkar, 2010, Can the production network be the testbed? In Proceedings of the 9th USENIX conference on Operating systems design and implementation (OSDI'10). USENIX Association, Berkeley, CA, USA, 1-6.

[Siat05] C. Siaterlis and V. Maglaris, "One step ahead to multisensor data fusion for DDoS detection," Journal of Computer Security, vol 13, Number 5, pp. 779-806, 2005

[Spri02] N. Spring, D. Wetherall, and T. Anderson, “Scriptroute: A public internet measurement facility,” in Proceedings of the 4th USENIX Symposium on Internet Technologies and Systems (USITS), 2002

[Stan02] S. Staniford, J.A. Hoagland, and J.M. McAlerney, Practical automated detection of stealthy portscans, in *Journal of Computer Security 10*, 2002, pp. 105 – 136

[Suur84] J. Suurballe and R. Tarjan, A Quick Method for Finding Shortest Pairs of Disjoint Paths, Networks, vol. 14, issue 2, 1984, pp. 325-336

[Tcpreplay] Tcpreplay , http://tcpreplay.appneta.com

[Teagle] Teagle, http://www.fire-teagle.org

[Teod09] P. García-Teodoro, J. Díaz-Verdejo, G. Maciá-Fernández, and E. Vázquez, Anomaly-based network intrusion detection: Techniques, systems and challenges, Computers & Security (Elsevier), vol. 28, no. 1-2, pp. 18 - 28, February-March 2009

[Vserver] Linux-VServer, http://linux-vserver.org

[vSwitch] VMware vSwitch, http://www.vmware.com/pdf/vi3_35/esx_3/r35u2/vi3_35_25_u2_3_server_config.pdf

[VXLAN] VXLAN: A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks, IETF Internet Draft – Experimental

[Wang08] Y. Wang, E. Keller, B. Biskeborn, J. van der Merwe, and J. Rexford, “Virtual Routers on the Move: Live Router Migration as a Network-management Primitive”, In Proc. SIGCOMM, August 2008

[Wu09] S.-Y. Wu, E. Yen, Data mining-based intrusion detectors, Expert Systems

with Applications, vol. 36, no. 3, pp. 5605 - 5612, April 2009

[XEN] The Xen® hypervisor, http://www.xen.org

[Yu08] M. Yu, Y. Yi, J. Rexford, and M. Chiang, Rethinking virtual network embedding: Substrate support for path splitting and migration, ACM SIGCOMM Computer Communication Review, vol. 38, no. 2, pp. 17–29, 2008

9 Δημοσιεύσεις στα πλαίσια της Διδακτορικής Διατριβής

9.1 Δημοσιεύσεις σε Διεθνή περιοδικά (με κρίση)

[J1] **C. Argyropoulos**, S. Mastorakis, G. Androulidakis, D. Kalogeras, V. Maglaris, "Network Virtualization in Multi-Tenant Software Defined Networking", Computer Networks, Elsevier (under review)

[J2] K. Giotis, **C. Argyropoulos**, G. Androulidakis, D. Kalogeras, V. Maglaris, "Combining OpenFlow and sFlow for an effective and scalable anomaly detection and mitigation mechanism on SDN environments", Computer Networks, Volume 62, 7 April 2014, Pages 122-136, ISSN 1389-1286, http://dx.doi.org/10.1016/j.bjp.2013.10.014

9.2 Δημοσιεύσεις σε Διεθνή Συνέδρια (με κρίση)

[C1] E. Jacob, A. Mendiola, L. Podleski, R. Krzywania, M. Przywecki, K. Dombek, A. Juszczyk, J. Aznar-Baranda, A. Vico-Oton, X. Jeannin, K. Baumann, **C. Argyropoulos**, "Multi-domain Software Defined Networking: Exploring possibilities", in Proc of the Terena Networking Conference (TNC2014), Dublin, Ireland, May 2014

[C2] **C. Argyropoulos**, G. Androulidakis, D. Kalogeras, B. Pietrzak, B. Belter, L. Lymberopoulos and V. Maglaris, "Network Virtualization over Heterogeneous Federated Infrastructures: Data Plane Connectivity", IFIP/IEEE Integrated Network Management Symposium (IEEE IM 2013), Ghent, Belgium, May 2013

[C3] **C. Argyropoulos**, D. Kalogeras, G. Androulidakis and V. Maglaris,, "PaFloMon -- A Slice Aware Passive Flow Monitoring Framework for OpenFlow Enabled Experimental Facilities," in 2012 European Workshop on Software Defined Networking (EWSDN), vol., no., pp.97,102, 25-26 Oct. 2012

[C4] C. Marinos, **C. Argyropoulos**, M. Grammatikou & V.Maglaris: "An Autonomic Monitoring Framework for QoS Management in Multi-Service Networks",in Proc. of the 2nd International ICST Conference on Mobile Networks & Management, Spain, Sept. 2010