



ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ

ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ & ΥΠΟΛΟΓΙΣΤΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

<<Σύγχρονες τάσεις οπτικής δικτύωσης για την εξυπηρέτηση εφαρμογών τεχνητής νοημοσύνης σε κέντρα δεδομένων>>

Γιαννακόπουλος Παναγιώτης

A.M 18390100

Εισηγητής: Μπόγρης Αντώνιος

Διπλωματική Εργασία

<<Σύγχρονες τάσεις οπτικής δικτύωσης για την εξυπηρέτηση εφαρμογών τεχνητής νοημοσύνης σε κέντρα δεδομένων>>

Εισηγητής:

<< ΑΝΤΩΝΙΟΣ , ΜΠΟΓΡΗΣ , ΚΑΘΗΓΗΤΗΣ >>

Εξεταστική Επιτροπή:

<< ΑΝΤΩΝΙΟΣ , ΜΠΟΓΡΗΣ , ΚΑΘΗΓΗΤΗΣ >>

<< ΝΙΚΟΛΑΟΣ , ΜΥΡΙΔΑΚΗΣ , ΑΝΑΠΛ. ΚΑΘΗΓΗΤΗΣ >>

<< ΠΑΝΑΓΙΩΤΗΣ , ΚΑΡΚΑΖΗΣ , ΑΝΑΠΛ. ΚΑΘΗΓΗΤΗΣ >>

Ημερομηνία εξέτασης : 9/10/2024

ΔΗΛΩΣΗ ΣΥΓΓΡΑΦΕΑ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ

Ο κάτωθι υπογεγραμμένος Γιαννακόπουλος Παναγιώτης του Χρήστου, με αριθμό μητρώου 18390100 φοιτητής του Τμήματος Μηχανικών Πληροφορικής και Υπολογιστών της Σχολής Μηχανικών του Πανεπιστημίου Δυτικής Αττικής, δηλώνω ότι:

«Βεβαιώνω ότι είμαι συγγραφέας αυτής της Διπλωματικής εργασίας και κάθε βοήθεια την οποία είχα για την προετοιμασία της, είναι πλήρως αναγνωρισμένη και αναφέρεται στην εργασία. Επίσης, οι όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών ή λέξεων, είτε ακριβώς είτε παραφρασμένες, αναφέρονται στο σύνολό τους, με πλήρη αναφορά στους συγγραφείς, τον εκδοτικό οίκο ή το περιοδικό, συμπεριλαμβανομένων και των πηγών που ενδεχομένως χρησιμοποιήθηκαν από το διαδίκτυο. Επίσης, βεβαιώνω ότι αυτή η εργασία έχει συγγραφεί από μένα αποκλειστικά και αποτελεί προϊόν πνευματικής ιδιοκτησίας τόσο δικής μου, όσο και του Ιδρύματος. Παράβαση της ανωτέρω ακαδημαϊκής μου ευθύνης αποτελεί ουσιώδη λόγο για την ανάκληση του πτυχίου μου».

Ο Δηλών



Γιαννακόπουλος Παναγιώτης

ΠΕΡΙΛΗΨΗ

Στο πλαίσιο αυτής της διπλωματικής εργασίας θα διερευνηθούν σύγχρονες τάσεις οπτικής δικτύωσης που χρησιμοποιούνται για την διακίνηση δεδομένων σε κέντρα δεδομένων. Κύριος σκοπός τους είναι η επίτευξη της χαμηλότερης καθυστέρησης κατά την μεταφορά δεδομένων, η ευελιξία σε συνδυασμό με την οικονομικά αποδοτικότερη υπηρεσία δεδομένων που θα προσφέρουν και η διάχυση όσο το δυνατόν μεγαλύτερης χωρητικότητας δεδομένων από τον έναν κόμβο στον άλλο. Κάποιες από αυτές τις σύγχρονες τάσεις οπτική δικτύωσης είναι η οπτική μεταγωγή, τα διαχωρισμένα υπολογιστικά συστήματα που ενεργοποιούνται βάση αυτής, όπως και οι ετερογενείς αρχιτεκτονικές υπολογιστών βασισμένες σε chiplet. Στην συγκεκριμένη έρευνα θα πραγματοποιηθεί μελέτη των παραπάνω τεχνολογιών σε εφαρμογές τεχνητής νοημοσύνης και μοντέλα μηχανικής μάθησης σε κέντρα δεδομένων.

ABSTRACT

In this thesis we will explore current optical networking trends used for data movement in data centers. Their main purpose is to achieve the lowest latency during data transfer, flexibility combined with the most cost-effective data service they offer and to diffuse as much data capacity as possible from one node to another. Some of these current optical networking trends are optical switching, separated computing systems enabled based on it, as well as heterogeneous chiplet-based computing architectures. In this research we will study these technologies in artificial intelligence applications and machine learning models in data centers.

Κατάλογος Σχημάτων

Σχήμα 1.1 : Μια αρθρωτή πλατφόρμα κέντρου δεδομένων που αναπτύχθηκε από το Τεχνολογικό Πανεπιστήμιο Nanyang (NTU) και την Toshiba και η αρχιτεκτονική δικτύου 2 επιπέδων που βασίζεται σε δέντρα.[1]

Σχήμα 1.2: Παραδείγματα τριών τυπικών αρχιτεκτονικών δικτύων με επίκεντρο τον μεταγωγέα(switch): Fat-Tree, VL2 και Jellyfish.[1]

Σχήμα 1.3: Παραδείγματα τριών τυπικών αρχιτεκτονικών δικτύων με επίκεντρο τον διακομιστή: DCell, BCube και FiConn.[1]

Σχήμα 1.4: Παραδείγματα τριών τυπικών υβριδικών αρχιτεκτονικών ηλεκτρονικής/οπτικής βάσης: Helios, HyPaC, OSA και Mordia.[1]

Σχήμα 1.5: Παραδείγματα και σύγκριση επιδόσεων δύο υβριδικών τοπολογιών βάσης ηλεκτρικών/ασύρματων στοιχείων[1]

Σχήμα 2.1 Εξαρτήματα ενός τυπικού οπτικού switch[2]

Σχήμα 2.2 Εξαρτήματα και λειτουργία ενός τυπικού μεταγωγέα πακέτων[2]

Σχήμα 2.3 Στοιχεία και λειτουργία ενός LION switch[2]

Σχήμα 2.4 Αρχιτεκτονική μεταγωγέα Retabit με $m + 1$ στάδια δρομολογητών AWGR[2]

Σχήμα 2.5 Στοιχεία και λειτουργία μιας Helios αρχιτεκτονικής[2]

Σχήμα 2.6 Αρχιτεκτονική και λειτουργία του EODCN[2]

Σχήμα 2.7 Τοπολογία μιας αρχιτεκτονικής FISSION[2]

Σχήμα 3.1 Δομή ενός CNN[7]

Σχήμα 3.2 Διάγραμμα μιας μονάδας RNN[7]

Σχήμα 3.3 Δομή ενός GAN[7]

Σχήμα 4.1 Αρχιτεκτονική ενός PINE συστήματος[5]

Σχήμα 4.2 Καθοδήγηση εύρους ζώνης στην ευέλικτη αρχιτεκτονική fat tree PINE[5]

Σχήμα 4.3 Μέση βελτίωση της απόδοσης του συστήματος του κατευθυνόμενου δέντρου σε σύγκριση με ένα vanilla fat tree χωρίς μείωση του εύρους ζώνης του ανώτερου στρώματος[5]

Σχήμα 4.4 Σχηματική απεικόνιση της βαθιάς διάσπασης και των προσαρμοσμένων διαμορφώσεων κόμβων για την υποστήριξη διαφορετικών φόρτων εργασίας[5]

Σχήμα 4.5 Φυσικά δίκτυα δοκιμών: (a) dragonfly, (b) fat tree και (c) implemented hardware[5]

Σχήμα 4.6 (α)-(γ) Προσεγγίσεις 2,5D και 3D MCM που διερευνήθηκαν στη φάση 1- δ) δίκτυο 8×8 στο τσιπ που υλοποιήθηκε στο ενεργό πλατφόρμα interposer.[5]

Σχήμα 4.7 (α) Πλήρως συσκευασμένος πρωτότυπος πομποδέκτης MCM (β) Πομποδέκτης MCM που χρησιμοποιεί παθητικό πυρίτιο interposer για ενσωμάτωση 2,5D.[5]

Σχήμα 4.8 (α) Φωτογραφία μικροσκοπίου της κατασκευασμένης διάταξης με ένθετα του διευρυμένου χωρικού πολυπλέκτη 4×1 με βάση το MRR, των διασταυρώσεων Si/SiN και του διαστρωματικού ζεύκτη. (β) Φωτογραφία μικροσκοπίου του βιδωμένου switch η 8×8 τριπλών στρωμάτων, (γ) Πακεταρισμένος switch με παρεμβαλλόμενο υλικό πυριτίου.[5]

Σχήμα 4.9 Προτεινόμενος σχεδιασμός switch χώρου και μήκους κύματος με χρήση συστοιχιών επιλογών μήκους κύματος με βάση το microring και συσσωρευτών χτενίσματος. Δείχνουν την αρχή λειτουργίας του πίνακα διασταυρούμενων σημείων $N \times M$ και πώς μπλοκάρει πλήρως τη διασταυρούμενη ομιλία πρώτης τάξης εντός της ζώνης.[5]

Σχήμα 4.10 (α) Σχηματική απεικόνιση του ανθεκτικού link plug-and-play. (β) εικόνα ηλεκτρονικής μικροσκοπίας σάρωσης των κατασκευασμένων συσκευών[5]

Σχήμα 4.11 (α) Σχηματικό διάγραμμα της διάταξης λεπτών ινών για οκτώ ίνες με βήμα 30 μm . (β) Σύζευξη συστοιχίας ινών σε συστοιχία οκτώ ζεύκτες plug-and-play γραμμένοι απευθείας στο τσιπ φωτονικής.[5]

Σχήμα 4.12 (α) Μοντέλα φωτονικής ζεύξης διαθέσιμα στο PhoenixSim εργαλείο λογισμικού. (β) Αρχιτεκτονική σύνδεσης DWDM εξαιρετικά χαμηλής κατανάλωσης με α πηγή λείζερ comb. Οι γραμμές comb DWDM αποσυμπλέκονται σε τέσσερις ομάδες, και κάθε ομάδα έχει διαδοχικούς διαμορφωτές που βασίζονται σε μικροδακτυλίους.[5]

Σχήμα 4.13 (α) Εικόνα μικροσκοπίου μιας ενσωματωμένης πηγής comb (αριστερά) και φωτογραφία μιας πλήρως ενσωματωμένης πηγής comb (δεξιά) (β) Φάσμα comb Soliton που δημιουργείται με χρήση μπαταρίας AAA παρέχοντας 98 mW ηλεκτρικής ενέργειας.[5]

Σχήμα 4.14 Σχέδια κατασκευών MZI και WDM με μειωμένη ευαισθησία στις παραλλαγές κατασκευής.[5]

Σχήμα 4.15 Μετάδοση (α) ενός παραδοσιακά σχεδιασμένου MZI που βασίζεται στα 500 nm με κυματοδηγό απλής λειτουργίας και (β) MZI με βάση προκαταρκτικά σχέδια που αποτελούνται από το πρόσφατα σχεδιασμένο πλάτος 1,2 μm οδηγός κύματος.[5]

Σχήμα 4.16 α) Αρχιτεκτονική Flex-SiPAC που βασίζεται στην αναδρομική τοπολογία BCube . (β) Παράδειγμα δικτύου Flex-SiPAC($r = 2$, $L = 2$) που εμφανίζει δύο συνδέσεις θύρας ανά CU και πλήρη συνδεσιμότητα στο επίπεδο 0.[6]

Σχήμα 4.17 Κλιμάκωση εύρους ζώνης Flex-SiPAC υπό ομοιόμορφη κατανομή των μηκών κύματος.[6]

Σχήμα 4.18 (α) Παράδειγμα αρχιτεκτονικής πομποδέκτη με δυνατότητα επαναδιαμόρφωσης μήκους κύματος 1×2 . (β) Ομοιόμορφη κατανομή των εισερχόμενων διαμορφωμένων μήκη κύματος στις δύο θύρες εξόδου.[6]

Κατάλογος Πινάκων

Πίνακας 1.1: Σύγκριση τεχνολογιών μετάδοσης του DCN[1]

Πίνακας 1.2: Μετρική για την σύγκριση τοπολογιών που βασίζονται σε ηλεκτρικά στοιχεία[1]

Πίνακας 1.3: Λειτουργίες ενός δικτύου κέντρου δεδομένων[1]

Πίνακας 1.4: Σύνοψη και σύγκριση σχετικά με την κατανομή του εύρους ζώνης[1]

Πίνακας 1.5:Μείωση ισχύος για διαφορετικά κουμπιά δικτύου[1]

Πίνακας 1.6:Ταξινόμηση και στατιστικά στοιχεία των αστοχιών σε DCNs[1]

Πίνακας 2.1: Σύγκριση των μηχανισμών αναζήτησης όσον αφορά τη μνήμη τους, απαιτήσεις σε πόρους επεξεργασίας, την καθυστέρηση και κατά πόσον η αρχική ρύθμιση απαιτείται[2]

Πίνακας 2.2 : Κατανάλωση ενέργειας, μέγιστη απόδοση και καθυστέρηση διαφορετικών επιλογών εξαρτημάτων[2]

Πίνακας 2.3 Σύγκριση αρχιτεκτονικών βάσει στοιχείων[2]

Πίνακας 4.1 Κατάλογος μαθηματικών συμβόλων με τις αντίστοιχες περιγραφές τους[6]

Περιεχόμενα

Κεφάλαιο 1 : Κέντρα δεδομένων και τα δίκτυά τους	14
1.1.1 Εισαγωγή στα κέντρα δεδομένων	14
1.1.2 Δίκτυο Κέντρου Δεδομένων	14
1.1.3 Περιπτώσεις για την βελτιστοποίηση του σχεδιασμού και της λειτουργίας των DCN.....	14
1.1.4 Προκλήσεις και απαιτήσεις του DCN.....	14
1.2.1 Υποδομή του DCN.....	15
1.2.2 Τεχνολογίες μετάδοσης	15
1.2.3 Επικοινωνία μεταξύ των DCNs	17
1.3.1 Λειτουργίες δικτύου κέντρου δεδομένων.....	25
1.3.2 Έλεγχος κυκλοφορίας σε δίκτυα κέντρων δεδομένων	26
1.3.3 Χρήση του Δικτύου.....	26
1.3.4 Διαμοιρασμός εύρους ζώνης(Bandwidth sharing).....	27
1.3.5 Καθυστέρηση υπηρεσίας	28
1.3.6 Κατανάλωση ενεργειακών πόρων	28
1.3.7 Συντήρηση δικτύου.....	29
Κεφάλαιο 2 : Οπτική μεταγωγή σε κέντρα δεδομένων	30
2.1 Γενική αρχιτεκτονική οπτικού μεταγωγέα.....	30
2.2.1 Επιλογές εξαρτημάτων Switch	31
2.2.2 Διεπαφή εισόδου-εξόδου	31
2.2.3 Δίκτυο μεταγωγής (Switch Fabric).....	32
2.2.4 Port buffers.....	34
2.3.1 Οπτικές αρχιτεκτονικές DCN	39
2.3.2 Αρχιτεκτονική LION.....	39
2.3.3 Αρχιτεκτονική Petabit	41
2.3.4 Αρχιτεκτονική Helios	42
2.3.5 Αρχιτεκτονική EODCN	43
2.3.6 Αρχιτεκτονική OGDCN.....	45
2.3.7 Αρχιτεκτονική OSA	45
2.3.8 FISSION	46

2.4 Σύγκριση αρχιτεκτονικών	47
Κεφάλαιο 3 : Βασικές αρχές τεχνητής νοημοσύνης και οι απαιτήσεις της	49
3.1 Εισαγωγή στην τεχνητή νοημοσύνη	49
3.2 Βασικές αρχές τεχνητής νοημοσύνης.....	49
3.3 Απαιτήσεις της Τεχνητής Νοημοσύνης	57
3.4.1 Συνελκτικά Νευρωνικά Δίκτυα(Convolutional Neural Networks – CNNs). 61	
3.4.2 Αρχιτεκτονική του CNN.....	62
3.4.3 Πεδία εφαρμογής των CNNs.....	64
3.5.1 Αναδρομικά Νευρωνικά Δίκτυα (Recurrent Neural Networks – RNNs).....	65
3.5.2 Παραγωγικό Αντιφατικό Δίκτυο (Generative Adversarial Networks – GANs)	66
3.5.3 Λειτουργία των GANs.....	67
3.5.4 Κατηγορίες GANs	68
3.5.4 Πεδία εφαρμογής των GANs.....	69
Κεφάλαιο 4 : Οπτικά δίκτυα για ΑΙ εφαρμογές.....	70
4.1 Εισαγωγή στα φωτονικά ολοκληρωμένα και ενεργειακά αποδοτικά δίκτυα στα κέντρα δεδομένων(PINE).....	70
4.2 Αρχιτεκτονική ενός συστήματος PINE	72
4.3 Οπτικά διασυνδεδεμένα Multi-Chip Modules (MCM)	81
4.4 Φωτονικές συνδέσεις πυριτίου DWDM	86
4.5 Αρχιτεκτονική Flex-SiPAC	90
4.6 Συμμετοχικός Σχεδιασμός Συλλογικού Αλγορίθμου	96
Συμπεράσματα	97
Βιβλιογραφία	98

Κεφάλαιο 1 : Κέντρα δεδομένων και τα δίκτυά τους

1.1.1 Εισαγωγή στα κέντρα δεδομένων

Τα κέντρα δεδομένων (Data Centers – DCs), λόγω της ραγδαίας αύξησης των υπηρεσιών στον παγκόσμιο ιστό αποτελούν μία αναγκαία και σημαντική υποδομή για την υποστήριξη αυτής της συνεχώς αυξανόμενης ζήτησης. Ένα DC συνήθως φιλοξενεί έναν μεγάλο αριθμό υπολογιστικών και αποθηκευτικών κόμβων, οι οποίοι είναι διασυνδεδεμένοι με ένα αποκλειστικά σχεδιασμένο δίκτυο επικοινωνίας για αυτούς, γνωστό και ως δίκτυο κέντρου δεδομένων(Data Centers Network – DCN).

1.1.2 Δίκτυο Κέντρου Δεδομένων

Το δίκτυο κέντρου δεδομένων(DCN) λειτουργεί ως ο σκελετός για την επικοινωνία μεταξύ των DCs και κατέχει κρίσιμο ρόλο όσον αφορά στην αποδοτική τους λειτουργία. Το DCN σε αντίθεση με τα παραδοσιακά δίκτυα, λόγω των ιδιαίτερων απαιτήσεών του όπως η ποικιλία εφαρμογών, η μεγάλη κλίμακα, η υψηλή πυκνότητα ισχύος και η υψηλή αξιοπιστία, δημιουργεί συνεχώς σημαντικές προκλήσεις για την υποδομή και τις λειτουργίες του.

1.1.3 Περιπτώσεις για την βελτιστοποίηση του σχεδιασμού και της λειτουργίας των DCN

Ανά τακτά χρονικά διαστήματα έχουν παρατηρηθεί αρκετές ερευνητικές προσπάθειες για τη βελτιστοποίηση του σχεδιασμού και των λειτουργιών των DCs. Αυτές οι ερευνητικές προσπάθειες μπορούν να ταξινομηθούν σε δύο κατηγορίες: 1) υποδομές DCN και 2) λειτουργίες DCN. Όσον αφορά για την πρώτη κατηγορία γίνεται σύγκριση των τεχνολογιών μετάδοσης με τις τοπολογίες δικτύου που χρησιμοποιούνται ή συστήνονται για τις υποδομές DCN. Για την δεύτερη κατηγορία συνοψίζονται οι ήδη υπάρχουσες τεχνικές διαχείρισης της κυκλοφορίας δεδομένων και οι μέθοδοι που έχουν υλοποιηθεί για να εκπληρώνουν διάφορους στόχους στον επιχειρησιακό τομέα ,όπως η υψηλή χρήση του δικτύου, η ισότιμη κατανομή του εύρους ζώνης, η χαμηλή καθυστέρηση των εφαρμογών, η υψηλή ανθεκτικότητα και η χαμηλή κατανάλωση ενεργειακών πόρων.

1.1.4 Προκλήσεις και απαιτήσεις του DCN

Συγκρίνοντας την δομή και την λειτουργική φιλοσοφία του DCN με άλλα παραδοσιακά δίκτυα όπως τα τοπικά δίκτυα και τα δίκτυα ευρείας ζώνης, παρατηρείται ένα σύνολο αναγκών που πρέπει να αντιμετωπιστούν. Παρακάτω παρατίθενται κάποιες προκλήσεις οι οποίες θεωρούνται θεμελιώδης για τον σχεδιασμό και την λειτουργία ενός DCN :

Μεγάλη κλίμακα : Στην σύγχρονη πραγματικότητα ένα κέντρο δεδομένων καλείται να περιέχει χιλιάδες εκατοντάδες διακομιστές, ενώ ταυτόχρονα η κλίματα των σύγχρονων κέντρων δεδομένων αυξάνεται με εκθετικό ρυθμό. Ήδη η Microsoft φιλοξενεί περισσότερους από 1

εκατομμύριο διακομιστές σε περισσότερα από 100 κέντρα δεδομένων παγκοσμίως. Έτσι γίνεται εύκολα αντιληπτό ότι μία τεράστια κλίμακα συστήματος αυξάνει περαιτέρω τις

προκλήσεις ως προς τον σχεδιασμό του δικτύου όσον αφορά την διασύνδεση, την ανθεκτικότητα και το κόστος.

Πληθώρα ποικιλία εφαρμογών : Τα σύγχρονα κέντρα δεδομένων φιλοξενούν διάφορες υπηρεσίες και εφαρμογές. Κάποιες από αυτές είναι διαδικτυακές εφαρμογές όπως για παράδειγμα το ηλεκτρονικό ταχυδρομείο αλλά και διάφορες αναζητήσεις στο διαδίκτυο. Επίσης τα σύγχρονα κέντρα δεδομένων φιλοξενούν και υποδομές υπηρεσιών όπως καταναλωμένα συστήματα αρχείων και καταναλωμένες μηχανές εκτέλεσης. Τέτοιες υπηρεσίες και εφαρμογές σε ένα κέντρο δεδομένων καθιστούν μία ποικιλία διαφορετικών χαρακτηριστικών μετάδοσης του δικτύου.

Υψηλή κατανάλωση ενέργειας : Σύμφωνα με έρευνες η ετήσια κατανάλωση ενέργειας στις ΗΠΑ εκτιμήθηκε ότι ξεπερνούσε τις 100 δισεκατομμύρια kWh το 2011 με αντίστοιχο ετήσιο κόστος ηλεκτρικής ενέργειας ύψους 7,4 δισεκατομμυρίων δολαρίων. Παρόλο που οι δικτυακές συσκευές καταναλώνουν περίπου μονάχα το 10% της ενέργειας από τον συνολικό αριθμό κατανάλωσης ενέργειας στον κλάδο της ΤΠΕ(Τεχνολογίας Πληροφοριών και επικοινωνιών),το ποσοστό αυτό μπορεί να αυξηθεί μέχρι και στο 50% στο μέλλον στην περίπτωση που οι διακομιστές γίνουν πιο ενεργειακά ανάλογοι με βάση το φορτίο εργασίας τους.

Αυστηρές απαιτήσεις υπηρεσιών : Μία θεμελιώδης απαίτηση για την φιλοξενία υπηρεσιών σε κέντρα δεδομένων είναι η συνεχής παροχής τους κάθε ώρα και κάθε μέρα ανά έτος, κάτι για οποίο απαιτείται η υψηλή ανθεκτικότητα του συστήματος. Ωστόσο, πιθανές βλάβες είτε στο υλικό είτε στο λογισμικό ή ακόμα και από λάθη τρίτων ατόμων μπορούν να είναι αναπόφευκτες. Για αυτό χρειάζεται η συνεχής παρακολούθηση του συστήματος και η άμεση και ευέλικτη αποκατάσταση των βλαβών.

1.2.1 Υποδομή του DCN

Η υποδομή ενός δικτύου κέντρου δεδομένων όπως ήδη έχει αναφερθεί διασυνδέει διακομιστές μέσα σε ένα κέντρο δεδομένων και μεταξύ άλλων κέντρων δεδομένων. Η

υποδομή αυτή μπορεί να κατηγοριοποιηθεί σε δύο διαστάσεις : i) την τεχνολογία μετάδοσης(ηλεκτρική, ασύρματη και οπτική) όπως και ii) την κλίμακα.

1.2.2 Τεχνολογίες μετάδοσης

Κάθε τεχνολογία μετάδοσης όπως είναι η ηλεκτρική, η ασύρματη και η οπτική χρησιμοποιούνται ευρέως σε υπολογιστικά δίκτυα. Αναλόγως το πεδίο εφαρμογής τους η κάθε τεχνολογία μετάδοσης παρουσιάζει και τα δικά της οφέλη, τα οποία διακρίνονται στον Πίνακα 1.1. Παραδοσιακά είναι συνηθής η χρήση των ηλεκτρικών καλωδίων στα δίκτυα των κέντρων δεδομένων(DCNs), ενώ παράλληλα ευρεία γίνεται και η χρήση των οπτικών ινών εξαιτίας της χαμηλής απώλειας απόδοσης, της χαμηλής κατανάλωσης ενέργειας αλλά και τον υψηλό ρυθμό της μετάδοσης των δεδομένων. Τέλος όσον αφορά την ασύρματη δικτύωση, επιλέγεται συνήθως όταν κρίνεται αναγκαία η ευελιξία στην διασύνδεση των διακομιστών.

Medium	Standard ¹	Bandwidth	Range ²	Applications
electrical cable	1000BASE-T, 10GBASE-T	1000 Mbps, 10 Gbps	100 m	intra DCN
optical fiber	40GBASE-LR4, 100GBASE-LR4, Infiniband	40 Gbps, 100 Gbps	10 km	intra and inter DCN
radio wave	IEEE 802.11.ad	7 Gbps	10 m	intra DCN

¹Only a few representative standards are listed.

²These are typical appropriate range in the standards.

Πίνακας 1.1: Σύγκριση τεχνολογιών μετάδοσης του DCN[1]

Παρακάτω αναλύονται οι κατηγορίες των τεχνολογιών μετάδοσης στο δίκτυο κέντρου δεδομένων όπως αναφέρθηκαν παραπάνω:

Ηλεκτρική : Χρησιμοποιούνται ηλεκτρικά καλώδια, συνήθως συνεστραμμένου ζεύγους(π.χ. καλώδια cat 5 και cat 6), λόγω των πλεονεκτημάτων που προσφέρουν στην εύκολη ανάπτυξη αλλά και την συμβατότητα με κάποιο παλαιότερο σύστημα δικτύου. Ένα συνεστραμμένο ζεύγος καλωδίων είναι ικανό να υποστηρίξει ταχύτητες των 10 Mbps, 100 Mbps, 1000 Mbps και των 10 Gbps κάθε μία με διαφορετικές προδιαγραφές. Παρόλα αυτά μπορούν να χρησιμοποιηθούν μονάχα εντός ενός κέντρου δεδομένων καθώς συνήθως επιβάλλουν ένα όριο μήκους 100 μέτρων χωρίς την χρήση αναμεταδοτών ή μεταγωγών.

Οπτική : Η συγκεκριμένη τεχνολογία μετάδοσης χρησιμοποιεί ως θεμελιώδη βάση τις οπτικές ίνες. Πρόκειται για μία τεχνολογία για τα κέντρα δεδομένων η οποία παρέχει υψηλό εύρος ζώνης, μικρότερος μέγεθος και βάρος και χαμηλότερη κατανάλωση ενέργειας σε σχέση με την ηλεκτρική τεχνολογία μετάδοσης. Γίνεται χρήση της τεχνικής της πολυπλεξίας διαίρεσης μήκους του κύματος(WDM) και αναπτύσσονται οπτικές συνδέσεις έως και 100 Gbps, ενώ αντίθετα οι ηλεκτρικές συνδέσεις περιορίζονται στα περίπου 40 Gbps. Οι οπτικές ίνες μπορούν να συνεργαστούν με ηλεκτρικούς switch με πομποδέκτες, όπως τον GBIC που είναι μετατροπέας διασύνδεσης των gigabits και τον SFP(small form-factor pluggable). Επιπλέον οι οπτικές ίνες μπορούν να κατευθύνουν μία δέσμη φωτός από οποιαδήποτε θύρα εισόδου σε οποιαδήποτε θύρα εξόδου, χωρίς λοιπόν να απαιτείται η χρήση κάποιας φωτοηλεκτρικής μετατροπής. Η οπτική τεχνολογία μετάδοσης βρίσκει πεδίο εφαρμογής σε διάφορα συστήματα όπως μικροηλεκτρομηχανικά συστήματα(MEMS), θερμο-οπτικά φαινόμενα, ηλεκτρο-οπτικά φαινόμενα κτλ.

Ασύρματη δικτύωση : Σε αυτήν την τεχνολογία μετάδοσης χρησιμοποιείται ασύρματη επικοινωνία που παρέχει τεχνολογία πολλαπλών gigabit ταχυτήτων που λειτουργούν στην μη αδειοδοτημένη συχνότητα των 60 GHz(π.χ. IEEE 802.11 ad) και προσφέρει μεγαλύτερο εύρος ζώνης. Σε ένα κέντρο δεδομένων όπου οι διακομιστές είναι συγκεντρωμένοι σε μικρές αποστάσεις μεταξύ τους και το ραδιοπεριβάλλον είναι στατικό, η χρήση των ασύρματων δικτύων θεωρείται ως ωφέλιμη. Εφόσον λοιπόν οι τυχόν παρεμβολές σήματος μπορούν σε ένα τέτοιο περιβάλλον να αποφευχθούν, γίνεται δυνατή η ανάπτυξη ασύρματων συσκευών ακολουθώντας την αρχιτεκτονική μεταγωγών top-of-racks.

1.2.3 Επικοινωνία μεταξύ των DCNs

Η επικοινωνία μεταξύ των δικτύων δεδομένων είναι εξαιρετικά πολύπλοκη, αφού συνδέει ένα τεράστιο πλήθος συσκευών με κρίσιμες απαιτήσεις απόδοσης, ενώ ταυτόχρονα παραμένει πλήρως ελεγχόμενη από έναν μοναδικό διαχειριστή. Και οι τρεις μορφές τεχνολογιών μετάδοσης έχουν μελετηθεί για χρήση σε εσωτερικά δίκτυα δεδομένων. Σε αυτή την ενότητα, εξετάζουμε τις υπάρχουσες τοπολογίες δικτύων εντός των κέντρων δεδομένων ανάλογα με τις τεχνολογίες μετάδοσης που εφαρμόζονται, δηλαδή, την τοπολογία με βάση τα ηλεκτρικά στοιχεία, την τοπολογία που βασίζεται σε ηλεκτρικά/οπτικά στοιχεία, και την τοπολογία που βασίζεται σε ηλεκτρικά/ασύρματα στοιχεία. Η τοπολογία που βασίζεται σε ηλεκτρικά στοιχεία εξακολουθεί να είναι η κυρίαρχη στη δικτύωση εντός των κέντρων δεδομένων. Η οπτική μετάδοση προβάλλεται ως η τεχνολογία του μέλλοντος για δικτύωση κέντρων δεδομένων με υψηλό εύρος ζώνης και μειωμένο κόστος. Παράλληλα, η ασύρματη μετάδοση θεωρείται ως λύση για την ενίσχυση της υπάρχουσας δικτυακής υποδομής χωρίς επιπλέον κόστος καλωδίωσης.

1.2.3.1 Τοπολογίες βασισμένες σε ηλεκτρικά στοιχεία

Τα περισσότερα εσωτερικά κέντρα δεδομένων βασίζονται σε ηλεκτρονικούς switch λόγω των πλεονεκτημάτων της ευκολίας εγκατάστασης και της συμβατότητας με τα παλαιότερα δικτυακά συστήματα. Ανάλογα με τον ρόλο των συσκευών, οι τοπολογίες που βασίζονται σε ηλεκτρικά στοιχεία είναι γενικά οργανωμένες σε διακοπιοκεντρική ή διακομιστοκεντρική μορφή. Οι στόχοι σχεδιασμού της τοπολογίας επικεντρώνονται συνήθως στο εύρος ζώνης διχοτόμησης, στην απόσταση άλματος, στην πολυπλοκότητα της καλωδίωσης, στην ανοχή σφαλμάτων, στο κόστος κατασκευής και στο αυξητικό κόστος αναβάθμισης.

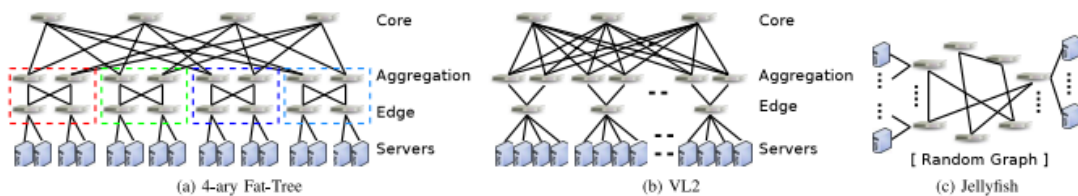
α) **Τοπολογίες με επίκεντρο τον switch:** Οι switch αναλαμβάνουν την πρωταρχική ευθύνη για τη δρομολόγηση και τη μετάδοση δεδομένων στο δίκτυο. Οι switch συνήθως συνδέονται σε ιεραρχικές τοπολογίες, ενώ οι διακομιστές συνδέονται γενικά με τους switch χαμηλού επιπέδου που βρίσκονται στην άκρη του δικτύου.

Τα δίκτυα στα συμβατικά κέντρα δεδομένων κατασκευάζονται συνήθως με βάση μια ιεραρχική τοπολογία που βασίζεται σε δέντρα με δύο ή τρία επίπεδα, ανάλογα με την κλίμακα του δικτύου. Για τα αρθρωτά κέντρα δεδομένων, όπως φαίνεται στο Σχήμα 1.1, μια αρχιτεκτονική δύο επιπέδων αποτελείται από μεταγωγείς ακραίας βαθμίδας και μεταγωγείς κεντρικής βαθμίδας για την υποστήριξη εκατοντάδων ή χιλιάδων διακομιστών. Για μεγάλης κλίμακας κέντρα δεδομένων, η αρχιτεκτονική τριών επιπέδων μπορεί να είναι πιο κατάλληλη, όπου οι μεταγωγείς είναι διατεταγμένοι σε τρία επίπεδα: άκρη, συνάθροιση και πυρήνα. Γενικά, όσο υψηλότερη είναι η βαθμίδα, τόσο οι μεταγωγείς είναι υψηλότερης ταχύτητας, αλλά και πιο ακριβοί και ενεργοβόροι. Ωστόσο, η υπερ-συνδρομή είναι συνηθισμένη. Για παράδειγμα, σε μια σύνδεση 1 Gbps, όταν το εύρος ζώνης της σύνδεσης μεταξύ άκρης και συνάθροισης είναι 10 Gbps, αν 20 διακομιστές είναι συνδεδεμένοι στον μεταγωγέα άκρης, η αναλογία υπερσυνδρομής είναι 1:2. Επιπλέον, η συμβατική αρχιτεκτονική τριών επιπέδων πάσχει από σημαντικά προβλήματα όπως το χαμηλό εύρος ζώνης διχοτόμησης, η κακή επεκτασιμότητα και τα σημεία μοναδικής αποτυχίας.



Σχήμα 1.1 : Μια αρθρωτή πλατφόρμα κέντρου δεδομένων που αναπτύχθηκε από το Τεχνολογικό Πανεπιστήμιο Nanyang (NTU) και την Toshiba και η αρχιτεκτονική δικτύου 2 επιπέδων που βασίζεται σε δέντρα.[1]

Το Fat-Tree είναι μια κλασική τοπολογία για την επικοινωνία μεταξύ δικτύων κέντρων δεδομένων, η οποία διασυνδέει ομοίτυπους Ethernet μεταγωγείς για πλήρες εύρος ζώνης διχοτόμησης από άκρο σε άκρο. Το Σχήμα 1.2α δείχνει ένα παράδειγμα ενός 4-αρικού Fat-Tree. Ένα k -αρικό Fat-Tree μπορεί να υποστηρίξει $k^3/4$ κεντρικούς υπολογιστές χρησιμοποιώντας $5k^2/4$ μεταγωγείς, καθένας από τους οποίους έχει k θύρες. Οι μεταγωγείς συνδέονται σε τοπολογία τριών επιπέδων, δηλαδή, ένα στρώμα πυρήνα, ένα στρώμα συνάθροισης και ένα στρώμα άκρων. Οι μεταγωγείς άκρων και συνάθροισης είναι διατεταγμένοι σε k rods, καθένα από τα οποία περιλαμβάνει $k/2$ μεταγωγείς άκρων και $k/2$ μεταγωγείς συνάθροισης. Κάθε μεταγωγέας άκρων συνδέεται απευθείας με $k/2$ κεντρικούς υπολογιστές και οι υπόλοιπες $k/2$ θύρες συνδέονται με τους μεταγωγείς συνάθροισης. Υπάρχουν $k^2/4$ μεταγωγείς πυρήνα, καθένας από τους οποίους έχει μία θύρα συνδεδεμένη σε κάθε rod. Το πλεονέκτημα του Fat-Tree είναι ότι όλοι οι μεταγωγείς είναι ομοίτυποι και μπορούν να χρησιμοποιηθούν φθηνά προϊόντα για όλους τους μεταγωγείς. Επιπλέον, υπάρχουν πολλαπλοί ίσοι δρόμοι κόστους μεταξύ οποιωνδήποτε δύο κεντρικών υπολογιστών, και μπορεί να επιτευχθεί θεωρητικά πλήρες εύρος ζώνης διχοτόμησης. Το μειονέκτημα του Fat-Tree είναι η υψηλή πολυπλοκότητα της καλωδίωσης. Ένα Fat-Tree 48-αρικό περιλαμβάνει 27.648 διακομιστές, 2.880 μεταγωγείς και 82.944 καλώδια.



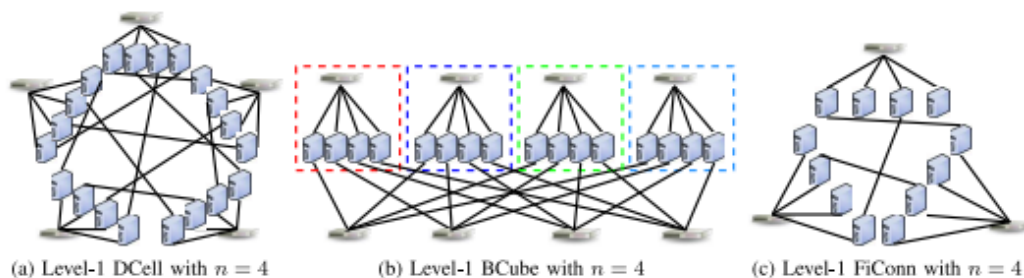
Σχήμα 1.2: Παραδείγματα τριών τυπικών αρχιτεκτονικών δικτύων με επίκεντρο τον μεταγωγέα (switch): Fat-Tree, VL2 και Jellyfish.[1]

Το VL2 χρησιμοποιεί επίσης μεταγωγείς χαμηλού επιπέδου για να σχηματίσει μια τοπολογία τριών επιπέδων βασισμένη σε δέντρο. Η διαφορά με το Fat-Tree είναι ότι η τοπολογία μεταξύ των μεταγωγέων πυρήνα και των μεταγωγέων συνάθροισης σχηματίζει ένα πλήρες διμερές γράφημα, και κάθε ακραίος μεταγωγέας συνδέεται με δύο μεταγωγείς συνάθροισης, όπως απεικονίζεται στο Σχήμα 2β. Το VL2 μειώνει τον αριθμό των καλωδίων αξιοποιώντας συνδέσεις μεταγωγέα-μεταγωγέα υψηλότερης ταχύτητας, π.χ. 10 Gbps για συνδέσεις μεταξύ διακοπών και 1 Gbps για συνδέσεις μεταξύ διακομιστών.

Το Jellyfish κατασκευάζει ένα τυχαίο κανονικό γράφημα στο επίπεδο ακρών, όπως απεικονίζεται στο Σχήμα 2γ. Συγκεκριμένα, μέρος των θυρών των ακραίων μεταγωγέων χρησιμοποιείται για τη σύνδεση διακομιστών και όλες οι υπόλοιπες θύρες χρησιμοποιούνται για να σχηματίσουν έναν τυχαίο γράφο των μεταγωγέων. Το Jellyfish επιτρέπει φυσικά τη σταδιακή επέκταση αφαιρώντας μια τυχαία υπάρχουσα σύνδεση και προσθέτοντας συνδέσεις από τον νέο μεταγωγέα στα δύο άκρα της αφαιρεθείσας σύνδεσης. Ένας αυθαίρετος διακομιστής στο Jellyfish μπορεί να φτάσει περισσότερους διακομιστές με λιγότερα άλματα σε σύγκριση με το Fat-Tree.

β) **Τοπολογίες με επίκεντρο τον διακομιστή:** Οι διακομιστές είναι απλώς τελικά σημεία του δικτύου. Ωστόσο, οι διακομιστές μπορούν να διαδραματίσουν σημαντικότερο ρόλο στο εσωτερικό δίκτυο του κέντρου δεδομένων, λόγω της ταχείας εξέλιξης του υλικού διακομιστών, όπως η πολυπύρνηνη CPU και οι πολλαπλές θύρες NIC. Για παράδειγμα, το GRIN αυξάνει τη χρήση του δικτύου βασισμένο σε μια τοπολογία VL2, διασυνδέοντας διακομιστές στο ίδιο ή σε γειτονικό rack χρησιμοποιώντας τις ελεύθερες θύρες τους.

Οι τοπολογίες με επίκεντρο τον διακομιστή αναδεικνύουν τη σημασία των διακομιστών στο δίκτυο ένα ακόμη βήμα μπροστά, θέτοντας τη διασύνδεση και τη νοημοσύνη δρομολόγησης στους διακομιστές και όχι στους μεταγωγείς. Στις τοπολογίες με επίκεντρο τον διακομιστή, οι διακομιστές δεν λειτουργούν μόνο ως τελικά σημεία αλλά και ως κόμβοι αναμετάδοσης μεταξύ τους. Η πιο συνήθης προσέγγιση για τοπολογίες με επίκεντρο τον διακομιστή είναι η αναδρομική κατασκευή ιεραρχικού δικτύου υψηλού επιπέδου από πολλαπλές μονάδες χαμηλού επιπέδου, όπως στην περίπτωση των DCell, BCube και FiConn.



Σχήμα 1.3: Παραδείγματα τριών τυπικών αρχιτεκτονικών δικτύων με επίκεντρο τον διακομιστή: DCell, BCube και FiConn.[1]

Το DCell έχει μια τυπική ιεραρχική τοπολογία με επίκεντρο τον διακομιστή, που επιτυγχάνεται μέσω μιας αναδρομικής κατασκευής. Σε ένα επίπεδο-0 DCell, n διακομιστές συνδέονται με έναν μεταγωγέα. Ένα DCell επιπέδου-1 κατασκευάζεται χρησιμοποιώντας $n + 1$ DCells επιπέδου-0. Συγκεκριμένα, μια θύρα κάθε διακομιστή σε κάθε DCell επιπέδου-0 συνδέεται με έναν διακομιστή σε άλλο DCell επιπέδου-0. Εάν οι DCells επιπέδου-0 θεωρηθούν ως εικονικοί κόμβοι, σχηματίζουν έναν πλήρη γράφο. Το Σχ. 3α δείχνει ένα παράδειγμα DCell επιπέδου-1 με $n = 4$. Το κυριότερο χαρακτηριστικό του DCell είναι η εξαιρετική επεκτασιμότητά του: ο αριθμός των υποστηριζόμενων διακομιστών διπλασιάζεται εκθετικά με τον αριθμό των θυρών NIC του διακομιστή. Ένα DCell επιπέδου-3 μπορεί να υποστηρίξει 3.263.442 διακομιστές με 4 θύρες στους διακομιστές και μεταγωγείς 6 θυρών.

Το BCube είναι μια τοπολογία που κατασκευάζεται αναδρομικά. Ένας BCube επιπέδου 0 αποτελείται από n διακομιστές συνδεδεμένους σε έναν μεταγωγέα n θυρών, ο οποίος είναι παρόμοιος με ένα DCell επιπέδου 0. Ένας BCube επιπέδου k ($k \geq 1$) κατασκευάζεται από n BCubes επιπέδου $(k-1)$ και nk μεταγωγείς n -θυρών, όπου κάθε μεταγωγέας συνδέεται με έναν διακομιστή σε κάθε BCube επιπέδου $(k-1)$. Το Σχήμα 3β δείχνει ένα παράδειγμα BCube επιπέδου-1 με αριθμό θυρών switch $n = 4$.

Από την άλλη, το FiConn πηγαίνει ένα βήμα παραπέρα στην κατασκευή της αρχιτεκτονικής περιορίζοντας τον αριθμό των θυρών NIC του διακομιστή σε 2. Συγκεκριμένα, ένα επίπεδο-0 FiConn αποτελείται από n διακομιστές και έναν μεταγωγέα n θυρών, όμοιο με το DCell επιπέδου-0 και το BCube επιπέδου-0. Ένα επίπεδο- k FiConn κατασκευάζεται χρησιμοποιώντας $b/2 + 1$ επίπεδο- $(k - 1)$ FiConns, όπου b είναι ο συνολικός αριθμός διακομιστών με διαθέσιμες θύρες σε κάθε επίπεδο- $(k - 1)$ FiConn. Τα επίπεδο- $(k - 1)$ FiConns σχηματίζουν ένα πλήρες γράφημα χρησιμοποιώντας $b/2$ θύρες από τις b διαθέσιμες εφεδρικές θύρες για να συνδεθούν μεταξύ τους. Στο Σχήμα 3γ παρουσιάζεται ένα παράδειγμα επιπέδου-1 FiConn με $n = 4$.

Αξίζει επίσης να αναφερθεί ότι, εκτός από τις ιεραρχικά δομημένες τοπολογίες, οι τοπολογίες με επίκεντρο τον διακομιστή μπορούν επίσης να δημιουργηθούν τυχαία. Για παράδειγμα, η Scafida προτείνεται για τη δημιουργία δικτύων με τυχαίο και ασύμμετρο τρόπο, εμπνευσμένη από τα δίκτυα χωρίς κλίμακα. Το Scafida κληρονομεί δύο κύριες ιδιότητες από τα δίκτυα χωρίς κλίμακα: εξαιρετικά μικρή διάμετρο και μεγάλη ανθεκτικότητα σε τυχαίες αποτυχίες. Ένα δίκτυο χωρίς κλίμακα μπορεί να δημιουργηθεί χρησιμοποιώντας την "προτιμησιακή προσάρτηση", δηλαδή ένας νέος κόμβος συνδέεται με έναν υπάρχοντα κόμβο με πιθανότητα ανάλογη του βαθμού του υπάρχοντος κόμβου. Η Scafida τροποποιεί τον αρχικό αλγόριθμο δημιουργίας δικτύου χωρίς κλίμακα περιορίζοντας τον μέγιστο βαθμό κόμβου, δεδομένου ότι οι διακομιστές και οι μεταγωγείς έχουν περιορισμένες θύρες. Τα αποτελέσματα προσομοίωσης δείχνουν ότι αυτή η τροποποίηση δεν επηρεάζει σημαντικά τις βασικές ιδιότητες των δικτύων χωρίς κλίμακα.

γ) Σύγκριση τοπολογιών βασισμένων σε ηλεκτρικά στοιχεία:

Οι τοπολογίες που βασίζονται σε ηλεκτρικά στοιχεία είναι οι πιο κοινές σε δίκτυα εντός κέντρων δεδομένων. Οι συγκρίσεις αυτών των τοπολογιών συχνά επικεντρώνονται σε λίγες βασικές πτυχές, όπως η επεκτασιμότητα, το κόστος κατασκευής, η χωρητικότητα, η

πολυπλοκότητα της καλωδίωσης, η λανθάνουσα κατάσταση και η ευρωστία. Ο αντίστοιχος δείκτης τους παρατίθεται στον Πίνακα 1.2.

Metric	Indicator
scalability	host number
construction cost	switch number, link number
capacity	bisection bandwidth ⁴
latency	diameter, hop distance
cabling complexity	link number
robustness	node degree, network connectivity under failures [36]

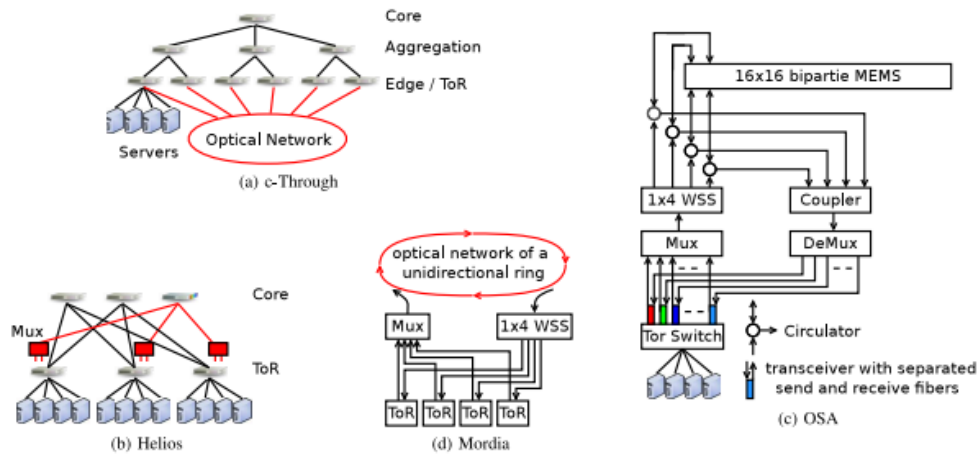
⁴Bisection bandwidth is the minimal bandwidth between the two equal parts of a network. It indicates network capacity under fully random traffic.

Πίνακας 1.2: Μετρική για την σύγκριση τοπολογιών που βασίζονται σε ηλεκτρικά στοιχεία[1]

Εκτός από τις τοπολογίες που δημιουργούνται με τη χρήση προκαθορισμένων διαδικασιών, οι τοπολογίες κέντρων δεδομένων μπορούν επίσης να δημιουργηθούν δυναμικά, προσαρμόζοντας τους στόχους βελτιστοποίησης. Το LEGUP εξάγει ένα λεπτομερές σχέδιο για το αναβαθμισμένο δίκτυο, επικεντρωμένο στη βελτιστοποίηση της ευελιξίας, της αξιοπιστίας και της ευκαμψίας, λαμβάνοντας υπόψη περιορισμούς όπως ο προϋπολογισμός, οι διαθέσιμοι τύποι μεταγωγέων και οι υπάρχουσες λεπτομέρειες των δικτύων κέντρων δεδομένων. Το LEGUP σχεδιάζει αποκλειστικά δενδροειδή δίκτυα που περιλαμβάνουν μεταγωγείς ακμής, συνάθροισης και πυρήνα. Γνωρίζοντας ότι οι τυχαίες τοπολογίες ενθαρρύνουν φυσικά τη σταδιακή αναβάθμιση των υφιστάμενων τοπολογιών, αρκετές ερευνητικές προσπάθειες έχουν επικεντρωθεί στη διερεύνηση της σκοπιμότητάς τους. Το REWIRE προσπαθεί να βρει δίκτυα με μέγιστο εύρος ζώνης διχοτόμησης και ελάχιστη καθυστέρηση από άκρο σε άκρο, ενώ παράλληλα ικανοποιεί περιορισμούς που καθορίζονται από τον χρήστη, όπως η τοπολογία και ο προϋπολογισμός. Το REWIRE σχεδιάζει την τοπολογία του δικτύου χρησιμοποιώντας μια τοπική αναζήτηση (π.χ., προσομοιωμένη απόπτηση) για να εξετάσει υποψήφιες τοπολογίες, δηλαδή τοπολογίες στις οποίες προστίθεται ή αφαιρείται ένας σύνδεσμος από τις υπάρχουσες τοπολογίες.

1.2.3.2 Τοπολογίες βασισμένες σε ηλεκτρικά/οπτικά στοιχεία

Με την ολοένα και αυξανόμενη ζήτηση για εύρος ζώνης στα κέντρα δεδομένων και τη συνεχώς μειούμενη τιμή των οπτικών συσκευών μεταγωγής, η οπτική μεταγωγή θεωρείται μια πολλά υποσχόμενη λύση για τα δίκτυα κέντρων δεδομένων. Ένα πρώτο βήμα για την εφαρμογή της οπτικής μεταγωγής στα κέντρα δεδομένων είναι ο συνδυασμός της συμβατικής ηλεκτρικής μεταγωγής με την οπτική μεταγωγή. Για παράδειγμα, η οπτική μεταγωγή μπορεί να χρησιμοποιείται για τη συγκέντρωση της κίνησης μεταξύ ραφιών(racks) ή μονάδων κέντρων δεδομένων (MDC). Δεδομένου ότι οι οπτικές συσκευές διαθέτουν ποικιλία χαρακτηριστικών και προδιαγραφών, οι τρέχουσες έρευνες για τα οπτικά δίκτυα σε κέντρα δεδομένων συχνά επικεντρώνονται στις συγκεκριμένες οπτικές συσκευές που εφαρμόζονται.



Σχήμα 1.4: Παραδείγματα τριών τυπικών υβριδικών αρχιτεκτονικών ηλεκτρονικής/οπτικής βάσης: Helios, HyPaC, OSA και Mordia.[1]

Το c-Through βασίζεται σε μια παραδοσιακή δενδρική τοπολογία, ενσωματώνοντας οπτικό δίκτυο που συνδέει ηλεκτρικούς switch ToR. Όπως φαίνεται στο Σχήμα 1.4α, το c-Through διαχειρίζεται την κυκλοφορία επεκτείνοντας τα όρια απομονωτικού χώρου υποδοχής TCP. Όταν το μέγεθος της απομονωμένης περιοχής υπερβαίνει ένα συγκεκριμένο όριο και η μεταβολή του μεγέθους της ουράς ξεπερνά ένα άλλο κατώφλι (και τα δύο κατώφλια είναι συνήθως 1 MB στην πράξη), τα στατιστικά στοιχεία αποστέλλονται σε έναν κεντρικό διαχειριστή. Ο διαχειριστής υπολογίζει την βέλτιστη διαμόρφωση της οπτικής διαδρομής, δηλαδή την καλύτερη αντιστοίχιση του πίνακα κίνησης μεταξύ των ρακοσυλλεκτών. Στη συνέχεια, ο διαχειριστής στέλνει εντολές αναδιάρθρωσης στο οπτικό δίκτυο και ειδοποιεί τον δαίμονα διαχείρισης του διακομιστή. Ο δαίμονας αναθέτει το αντίστοιχο αναγνωριστικό VLAN στα πακέτα, καθώς το ηλεκτρικό δίκτυο μεταγωγής και το οπτικό δίκτυο μεταγωγής ανήκουν σε διαφορετικά VLAN.

Το Helios παρουσιάζεται ως ένα πολυεπίπεδο δέντρο με πολλές ρίζες, περιλαμβάνοντας rod μεταγωγείς και μεταγωγείς πυρήνα, όπως φαίνεται στο Σχήμα 1.4β. Οι switch πυρήνα συνδυάζουν ηλεκτρικούς και οπτικούς switch. Παθητικοί οπτικοί πολυπλέκτες (που σημειώνονται ως "Mux") χρησιμοποιούνται για τη μεταφορά οπτικών σημάτων με διαφορετικά μήκη κύματος σε μια μόνο οπτική ίνα. Το Helios εκτιμά τη ζήτηση εύρους ζώνης χρησιμοποιώντας τη μέθοδο δικαιοσύνης max-min μεταξύ των ροών κυκλοφορίας, όπως μετράται στους switch rod. Παρόμοια με το c-Through, το Helios υπολογίζει τη διαμόρφωση της οπτικής διαδρομής βασισμένο σε αντιστοίχιση μέγιστης βαρύτητας. Ωστόσο, σε αντίθεση με το c-Through, στο Helios διαμορφώνονται δυναμικά τόσο οι ηλεκτρικοί όσο και οι οπτικοί switch σύμφωνα με την υπολογισμένη διαμόρφωση. Το Helios αποφασίζει πού θα προωθήσει την κυκλοφορία, είτε στο ηλεκτρικό δίκτυο είτε στο οπτικό δίκτυο. Το πρωτότυπο του Helios

δείχνει ότι η αλλαγή της διαδρομής κυκλοφορίας διαρκεί περίπου 266 ms συνολικά, εκ των οποίων 12 ms αφορούν το χρόνο οπτικής μεταγωγής.

Η OSA (Optical Switching Architecture) εξετάζει τη δυνατότητα δημιουργίας ενός πλήρως οπτικού δικτύου πυρήνα μεταξύ των μεταγωγέων ToR (Top-of-Rack). Όπως απεικονίζεται στο Σχήμα 4γ, οι οπτικοί πομποδέκτες που συνδέονται σε έναν μεταγωγέα ToR χρησιμοποιούν ξεχωριστές ίνες για αποστολή και λήψη.

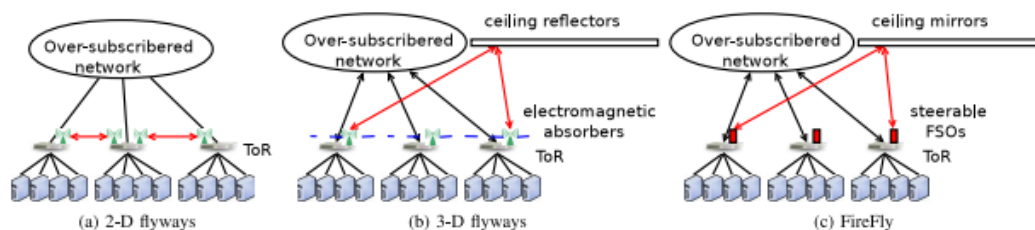
Η αρχιτεκτονική περιλαμβάνει διάφορες συσκευές όπως:

- 1) **Πολυπλέκτες (Mux)**: Συνδυάζουν οπτικά σήματα από πολλές ίνες σε μία μόνο ίνα.
- 2) **1×4 WSS (Wavelength Selective Switch)**: Προωθεί το οπτικό σήμα σε 4 θύρες, ανάλογα με το μήκος κύματος.
- 3) **Κυκλοφορητές**: Επιτρέπουν την αμφίδρομη μετάδοση σήματος.
- 4) **Διμερή οπτικό switch 16×16 MEMS**: Επιτρέπει συνδέσεις από οποιαδήποτε θύρα εισόδου σε οποιαδήποτε θύρα εξόδου.
- 5) **Συζεύκτες και αποπολυπλέκτες (DeMux)**: Συνδυάζουν σήματα με διαφορετικά μήκη κύματος σε μία ίνα και διαχωρίζουν ένα οπτικό σήμα με πολλά μήκη κύματος σε διαφορετικές ίνες.

Με βάση την εκτιμώμενη ζήτηση κυκλοφορίας, οι switch WSS και MEMS αναδιαμορφώνονται. Οι switch WSS και MEMS που χρησιμοποιούνται στην OSA επιβάλλουν χρόνους μεταγωγής 14 ms και 9 ms αντίστοιχα. Ο αντίκτυπος αυτών των χρόνων στις ροές κυκλοφορίας που είναι ευαίσθητες στην καθυστέρηση απαιτεί περαιτέρω διερεύνηση.

1.2..3.3 Τοπολογίες με βάση ηλεκτρικά/ασύρματα στοιχεία

Οι τρέχουσες μελέτες για τοπολογίες που βασίζονται σε ηλεκτρονικά και ασύρματα στοιχεία επικεντρώνονται στην αξιολόγηση της χρησιμότητας των ραδιοσυχνοτήτων των 60 GHz. Εξετάζουν επίσης την προσθήκη επιπλέον εύρους ζώνης επικοινωνίας σε παραδοσιακά δίκτυα που βασίζονται σε ηλεκτρικά στοιχεία.



Σχήμα 1. 5: Παραδείγματα και σύγκριση επιδόσεων δύο υβριδικών τοπολογιών βάσης ηλεκτρικών/ασύρματων στοιχείων[1]

Αρκετές μελέτες ήδη εξετάζουν την απόδοση των συνδέσεων 60 GHz σε περιβάλλον κέντρου δεδομένων. Τα αποτελέσματα δείχνουν ότι πολλαπλές συνδέσεις 60 GHz μπορούν να προσφέρουν την επιθυμητή απόδοση όσον αφορά τη διάδοση του σήματος, τη σταθερότητα της σύνδεσης και την επαναχρησιμοποίηση του διαθέσιμου χώρου. Ως αποτέλεσμα, προγραμματίζεται η δημιουργία μιας υβριδικής αρχιτεκτονικής δικτύου, η οποία θα ενσωματώνει ασύρματες συνδέσεις 60 GHz, γνωστές και ως "flyways", στην παραδοσιακή ηλεκτρονική αρχιτεκτονική για να ενισχύσει τη συνολική χωρητικότητα. Όπως φαίνεται στο Σχήμα 1.5α, κάθε μεταγωγέας ToR θα εξοπλιστεί με μία ή περισσότερες συνδέσεις 60 GHz που διαθέτουν κατευθυντικές κεραίες.

Με σκοπό την επιπρόσθετη βελτίωση της χρήσης των συχνοτήτων 60 GHz σε κέντρα δεδομένων, προτείνεται μια τρισδιάστατη λύση, όπως απεικονίζεται στο Σχήμα 1.5β, όπου ασύρματες συσκευές με περιστρεφόμενες κατευθυντικές κεραίες τοποθετούνται στην κορυφή των ραφιών (top-of-rack). Ανακλαστήρες οροφής, δηλαδή επίπεδες μεταλλικές πλάκες, χρησιμοποιούνται για την αντανάκλαση των σημάτων, ενώ ηλεκτρομαγνητικοί απορροφητές τοποθετούνται κοντά στις κεραίες για την αποφυγή τοπικών ανακλάσεων και σκέδασης. Αυτή η τρισδιάστατη διάταξη μειώνει την παρεμβολή, αποφεύγει τα εμπόδια και παρέχει μια έμμεση οπτική επαφή για αξιόπιστη επικοινωνία.

Με παρόμοιο σκεπτικό με τη τρισδιάστατη διάταξη δέσμης, το FireFly αξιοποιεί την τεχνολογία οπτικής ελεύθερης ζώνης (FSO) για να προσφέρει επιπλέον αναδιαμορφώσιμο εύρος ζώνης στο δίκτυο κέντρου δεδομένων. Σε σύγκριση με τις συνδέσεις στα 60 GHz, η τεχνολογία FSO παρέχει χαμηλότερες παρεμβολές και υψηλότερο εύρος ζώνης σε μεγαλύτερες αποστάσεις. Όπως δείχνεται στο Σχήμα 1.5γ, το FireFly χρησιμοποιεί καθρέφτες οροφής για την αντανάκλαση ακτίνων λέιζερ. Η κατεύθυνση των ακτίνων λέιζερ μπορεί να ρυθμιστεί με τη χρήση διακοπόμενων κατόπτρων ή κατόπτρων Galvo, επιτρέποντας έτσι τη δέσμη λέιζερ να κατευθυνθεί προς τον επιθυμητό στόχο.

1.2.3.4 Επικοινωνία εντός των κέντρων δεδομένων(Intra Data Centers)

Ένας μεγάλος αριθμός γεωγραφικά κατανομημένων κέντρων δεδομένων έχει αναπτυχθεί ως υποδομή για την παροχή υπηρεσιών με έμφαση στη χαμηλή καθυστέρηση και την αξιοπιστία. Από τη μία πλευρά, οι υπηρεσίες από ένα τοπικό κέντρο δεδομένων προσφέρουν συνήθως χαμηλή καθυστέρηση στους τελικούς χρήστες. Από την άλλη πλευρά, η δημιουργία αντιγράφων ασφαλείας και η ανάκτηση δεδομένων (συχνά υπό τη μορφή εικονικών μηχανών) σε γεωγραφικά κατανομημένα κέντρα δεδομένων μπορεί να αποτρέψει το ενιαίο σημείο αποτυχίας σε περιπτώσεις καταστροφών ή επιθέσεων.

Οι τοπολογίες των συγκεκριμένων δικτύων επηρεάζονται σε μεγάλο βαθμό από την επιλογή της θέσης των κέντρων δεδομένων και την ποιότητα της δικτυακής τους σύνδεσης. Η επιλογή των τοποθεσιών για τα κέντρα δεδομένων καθορίζεται από διάφορους παράγοντες:

1) **Γεωγραφική θέση:** Περιοχές με υψηλό κίνδυνο φυσικών καταστροφών, όπως σεισμοί, πλημμύρες, πυρκαγιές και τυφώνες, θα πρέπει να αποφεύγονται ή τα κέντρα δεδομένων σε αυτές τις περιοχές πρέπει να είναι σχεδιασμένα για να αντέχουν σε τέτοια φαινόμενα. Το κλίμα που επιτρέπει φυσική ψύξη, όπως δροσερός αέρας, αποτελεί επιπλέον πλεονέκτημα.

2) **Ηλεκτρική ενέργεια:** Η ηλεκτρική ενέργεια είναι κρίσιμη για τη λειτουργία των κέντρων δεδομένων. Το κόστος, η αξιοπιστία και η καθαρότητα της ενέργειας είναι σημαντικοί παράγοντες.

3) **Συνδεσιμότητα:** Υψηλής ποιότητας δικτυακή συνδεσιμότητα είναι απαραίτητη για την παροχή υπηρεσιών μέσω του δικτύου από το κέντρο δεδομένων.

4) **Επιχειρηματικό περιβάλλον:** Κανονισμοί φιλικό προς τις επιχειρήσεις και οικονομικά κίνητρα μπορούν να προσελκύσουν ιδιοκτήτες κέντρων δεδομένων. Προτιμώνται τοποθεσίες κοντά σε μεγάλες αγορές και πελάτες.

1.3.1 Λειτουργίες δικτύου κέντρου δεδομένων

Η λειτουργία του δικτύου του κέντρου διασφαλίζει τη μεταφορά δεδομένων από τις πηγές στους προορισμούς με ποικίλους σκοπούς, όπως η εγγύηση εύρους ζώνης, η αναλογική κατανομή εύρους ζώνης, η ισορροπία φορτίου και η ενεργειακή αποδοτικότητα. Σε αυτή την ενότητα, αρχικά εξετάζουμε τις τεχνικές ελέγχου της κυκλοφορίας που μπορούν να επηρεάσουν τον τρόπο μεταφοράς των δεδομένων, εστιάζοντας σε θέματα όπως η επιλογή διαδρομής, ο περιορισμός του ρυθμού και η διαχείριση προτεραιοτήτων. Ακολούθως, παρουσιάζουμε προτάσεις που στοχεύουν στην επίτευξη συγκεκριμένων στόχων, όπως αναφέρονται στον πίνακα 1.3.

DATA CENTER OPERATIONS				
Objective	Proposal	Traffic Control		
		Path Selection	Rate Control	Priority Management
Network Utilization	Hedera [42]	✓		
	B4 [60]	✓	✓	
	SWAN [62]	✓	✓	✓
	Seawall [63]		✓	
	SecondNet [64]	✓	✓	
Bandwidth Sharing	BwE [65]		✓	
	EyeQ [66], [67]		✓	
	D ³ [68]		✓	
	PDQ [69]		✓	
	pFabric [70]		✓	✓
Energy Consumption	D ² -TCP [71]		✓	
	GreenTE [72]	✓		
Network Maintenance	JHNOE [73]	✓		
	zUpdate [74]	✓		

Πίνακας 1.3: Λειτουργίες ενός δικτύου κέντρου δεδομένων[1]

1.3.2 Έλεγχος κυκλοφορίας σε δίκτυα κέντρων δεδομένων

Η βασική λειτουργία του δικτύου του κέντρου δεδομένων είναι να κατευθύνει τη ροή των δεδομένων από τις πηγές στους προορισμούς τους. Παραδοσιακά, η δρομολόγηση της κυκλοφορίας πραγματοποιείται με έναν κατανεμημένο τρόπο, όπου κάθε μεταγωγέας μαθαίνει την τοπολογία του δικτύου μέσω ανταλλαγής μηνυμάτων και δημιουργεί έναν πίνακα δρομολόγησης για την προώθηση των πακέτων. Σε αυτό το υποκεφάλαιο, θα εξεταστούν οι τεχνικές ελέγχου της κυκλοφορίας που μπορούν να επηρεάσουν αυτή τη διαδικασία από τρεις οπτικές γωνίες: την επιλογή διαδρομής, τον περιορισμό του ρυθμού και τη διαχείριση των προτεραιοτήτων.

1) **Επιλογή διαδρομής:** Οι διαδρομές μέσω των οποίων προωθούνται τα πακέτα στα δίκτυα DCN καθορίζονται από διάφορα πρωτόκολλα, όπως το spanning tree, τα VLANs, οι αλγόριθμοι δρομολόγησης, η δρομολόγηση πολλαπλών διαδρομών, η κωδικοποίηση πληροφοριών διαδρομής στα πακέτα, καθώς και το σύγχρονο πρωτόκολλο OpenFlow.

2) **Έλεγχος ρυθμού:** Ο έλεγχος ρυθμού είναι κρίσιμος για τη διαχείριση της συμφόρησης, την εξισορρόπηση του φορτίου και την παροχή εγγυημένου εύρους ζώνης σε ένα δίκτυο. Μπορεί να εφαρμοστεί τόσο στους τελικούς υπολογιστές όσο και στο δίκτυο.

3) **Διαχείριση προτεραιοτήτων:** Χωρίς διαχείριση προτεραιοτήτων, διαδραστικές ροές που έχουν μικρό όγκο δεδομένων αλλά απαιτούν γρήγορη απόκριση, συχνά καθυστερούν πίσω από μεγαλύτερες ροές. Η διαχείριση προτεραιοτήτων παρέχει διαφοροποιημένη ποιότητα υπηρεσίας, αντιμετωπίζοντας τα πακέτα βάσει της προτεραιότητάς τους αντί για τη σειρά άφιξής τους. Όταν ένα πακέτο φτάνει σε μια γεμάτη ουρά αποστολής ή λήψης, αν η προτεραιότητά του είναι ίση ή χαμηλότερη από την χαμηλότερη προτεραιότητα στο buffer, απορρίπτεται. Αντίθετα, το πακέτο με την χαμηλότερη προτεραιότητα απορρίπτεται για να δημιουργηθεί χώρος για το νέο πακέτο. Όταν μια θύρα είναι αδρανής, το πακέτο με την υψηλότερη προτεραιότητα επιλέγεται από την ουρά και αποστέλλεται.

1.3.3 Χρήση του Δικτύου

Λαμβάνοντας υπόψη την επένδυση κεφαλαίου για τις συσκευές δικτύου και το συνεχές κόστος για το εύρος ζώνης και τις λειτουργίες του δικτύου, είναι κρίσιμο να αξιοποιείται στο μέγιστο το διαθέσιμο εύρος ζώνης. Παρόλα αυτά, τα πραγματικά δεδομένα δείχνουν χαμηλή εκμετάλλευση των συνδέσεων τόσο στα εσωτερικά δίκτυα των κέντρων δεδομένων όσο και στα δίκτυα μεταξύ διαφορετικών κέντρων δεδομένων. Μία γενική προσέγγιση για την επίτευξη υψηλής απόδοσης του δικτύου είναι η εύρεση της ιδανικής διαδρομής, η κατανομή εύρους ζώνης στις ροές δεδομένων, και η εφαρμογή αυτής της κατανομής μέσω διαφόρων μεθόδων ελέγχου της κυκλοφορίας, όπως επιλογή διαδρομής, περιορισμός ρυθμού και διαχείριση προτεραιοτήτων.

1.3.4 Διαμοιρασμός εύρους ζώνης(Bandwidth sharing)

Παρόλο που τα σύγχρονα κέντρα δεδομένων μπορούν να προσφέρουν αποδοτικές κατανομές πόρων, όπως CPU, δίσκο και μνήμη, μέσω εικονικών μηχανών (VMs) ή κοντέινερς, το εύρος ζώνης του δικτύου εξακολουθεί να κατανέμεται με βέλτιστο αλλά όχι εγγυημένο τρόπο. Κακόβουλοι ενοικιαστές μπορούν να εκμεταλλευτούν αυτό το γεγονός για να βελτιώσουν άδικα την απόδοση του δικτύου τους, δημιουργώντας πολλαπλές συνδέσεις TCP ή χρησιμοποιώντας το πρωτόκολλο UDP. Αυτή η αδυναμία στον διαμοιρασμό εύρους ζώνης καθιστά δύσκολο τόσο για τους μισθωτές όσο και για τους παρόχους υπηρεσιών να δικαιολογήσουν την κατανομή των δικτυακών πόρων. Οι προσπάθειες για την εξασφάλιση δίκαιου διαμοιρασμού του εύρους ζώνης συνήθως επικεντρώνονται σε δύο βασικούς τομείς: την ελάχιστη εγγύηση εύρους ζώνης και την αναλογικότητα του εύρους ζώνης, ανάλογα με τα διαφορετικά συστήματα πληρωμών. Οι συγκρίσεις αυτών των προσεγγίσεων παρουσιάζονται αναλυτικά στον πίνακα 1.4.

SUMMARY AND COMPARISON ON BANDWIDTH ALLOCATION

Proposals	Objective	Hardware	Architecture	Control	Evaluation	Approach
Secondnet [64]	Bandwidth Guarantee	MPLS-based Switch	Arbitrarily	Centralized	Simulation & Testbed	static bandwidth reservation
Oktopus [104]	Bandwidth Guarantee	Commodity Switch	Arbitrarily	Centralized	Simulation & Testbed	admission control and static bandwidth reservation
Proteus [88]	Bandwidth Guarantee	Commodity Switch	Arbitrarily	Centralized	Simulation & Testbed	bandwidth reservation based on the time-varying requirement
EyeQ [66], [67]	Bandwidth Guarantee	Switch with ECN	non-blocking core	Distributed	Testbed	Min-guarantee via rate detection, control, and AC at end-host
ElasticSwitch [105]	Bandwidth Guarantee	Commodity Switch	Arbitrarily	Distributed	Testbed	hose-based minimum guarantee and dynamic rate allocation
Seawall [63]	Bandwidth Proportionality	Commodity Switch	Arbitrarily	Distributed	Testbed	weight-based proportion allocation
FairCloud [106]	Bandwidth Proportionality	One queue per-tenant (or VM)	Arbitrarily	Distributed	Simulation & Testbed	bandwidth allocation in proportion to the number of VMs of a tenant

Πίνακας 1.4: Σύνοψη και σύγκριση σχετικά με την κατανομή του εύρους ζώνης[1]

1) **Εγγύηση ελάχιστου εύρους ζώνης:** Συμφωνίες επιπέδου υπηρεσιών (SLAs): Όταν επιβάλλονται ρητές SLAs για το εύρος ζώνης, η ελάχιστη εγγύηση εύρους ζώνης διασφαλίζει ότι ο μισθωτής θα λάβει το ποσό του εύρους ζώνης για το οποίο έχει πληρώσει. Μια συνήθης μέθοδος για την παροχή αυτής της εγγύησης είναι η κράτηση εύρους ζώνης. Οι απαιτήσεις για το εύρος ζώνης μπορούν να καθοριστούν με διάφορους τρόπους, καθένας από τους οποίους απαιτεί διαφορετική υλοποίηση.

2) **Αναλογικότητα εύρους ζώνης:** Χωρίς την ύπαρξη επιπλέον SLAs για το εύρος ζώνης, η αναλογικότητα του εύρους ζώνης εξασφαλίζει ότι η διαθέσιμη ποσότητα πόρων για έναν μισθωτή είναι ανάλογη με το ποσό που έχει πληρώσει για άλλους πόρους, όπως CPU, μνήμη και δίσκο. Υπάρχουν διάφορες υλοποιήσεις της αναλογικότητας του εύρους ζώνης, όπως το Seawall και το FairCloud. Μια εναλλακτική προσέγγιση είναι η επέκταση των υφιστάμενων συστημάτων πληρωμών για να περιλαμβάνουν ρητά SLAs για το εύρος ζώνης, προσφέροντας διαφορετικές τιμές ανάλογα με το εγγυημένο εύρος ζώνης.

1.3.5 Καθυστέρηση υπηρεσίας

Η καθυστέρηση της υπηρεσίας είναι κρίσιμη για πολλές διαδραστικές εφαρμογές, όπως η αναζήτηση, τα κοινωνικά δίκτυα και τα διαδικτυακά παιχνίδια, οι οποίες παράγουν μικρές ποσότητες δεδομένων αλλά απαιτούν χαμηλό χρόνο απόκρισης. Ο χρόνος ολοκλήρωσης μιας ροής μπορεί να μειωθεί είτε με σιωπηρές μεθόδους, όπως η βελτιστοποίηση της κυκλοφορίας για υψηλή αποδοτικότητα του δικτύου, είτε με ρητό έλεγχο της κυκλοφορίας που βασίζεται σε προθεσμίες ή εκτιμήσεις χρόνου ολοκλήρωσης. Επιπλέον, η προτεραιοποίηση ροών ευαίσθητων στην καθυστέρηση μπορεί να μειώσει την απώλεια πακέτων και να βελτιώσει την καθυστέρηση κατά τη διάρκεια περιόδων συμφόρησης. Ο ρητός έλεγχος κυκλοφορίας απαιτεί πληροφορίες για τις προθεσμίες ή τα μεγέθη των ροών, οι οποίες μπορούν να ληφθούν από ανώτερα στρώματα, όπως εφαρμογές που δημιουργούν τις ροές. Με αυτές τις πληροφορίες, υπάρχουν θεωρητικά βέλτιστες στρατηγικές για τον έλεγχο της κυκλοφορίας. Ο αλγόριθμος SRTF (Shortest Remaining Time First) είναι γνωστός ως ο βέλτιστος για την ελαχιστοποίηση του μέσου χρόνου ολοκλήρωσης (FCT) σε μια ενιαία ζεύξη, επιλέγοντας προς αποστολή τις ροές με τα λιγότερα εναπομείναντα πακέτα πρώτες. Ο αλγόριθμος EDF (Earliest Deadline First) είναι ο βέλτιστος για τη μείωση της μέγιστης καθυστέρησης.

1.3.6 Κατανάλωση ενεργειακών πόρων

Η αυξημένη κατανάλωση ενέργειας αποτελεί πλέον ένα από τα πιο σημαντικά προβλήματα για τα σύγχρονα κέντρα δεδομένων, λόγω τόσο οικονομικών όσο και περιβαλλοντικών ανησυχιών. Πολλές προτάσεις εξοικονόμησης ενέργειας για τα δίκτυα των κέντρων δεδομένων έχουν αναδειχθεί στην πρόσφατη βιβλιογραφία και μπορούν να κατηγοριοποιηθούν σε δύο κύριες ομάδες: βελτιστοποίηση εντός του κέντρου δεδομένων και βελτιστοποίηση μεταξύ διαφορετικών κέντρων δεδομένων. Η πιο κοινή μέθοδος για την εξοικονόμηση ενέργειας στα κέντρα δεδομένων είναι η απενεργοποίηση ή η θέση σε κατάσταση ύπνου των αδρανών εξαρτημάτων, όπως οι συνδέσεις, οι θύρες και οι switch. Η βελτιστοποίηση μεταξύ των κέντρων δεδομένων αξιοποιεί τις τοπικές διαφορές στα χαρακτηριστικά των κέντρων δεδομένων, όπως το κόστος ενέργειας, την ενεργειακή απόδοση και τη χρήση καθαρής ενέργειας.

Knobs	Operation	Power saving
port	disable	up to 0.5 W/port
link rate	lower	0.4–1 W/port when a 1 Gb/s port is set to either 10 or 100 Mb/s
line card	power-off	30–60 W/card
switch	power-off	100–2000 W/switch for different device

Πίνακας 1. 5:Μείωση ισχύος για διαφορετικά κουμπιά δικτύου[1]

1) **Βελτιστοποίηση εντός του κέντρου δεδομένων:** Οι ενεργειακές βελτιστοποιήσεις στην πλευρά του δικτύου συνήθως βασίζονται στην απενεργοποίηση ή στην εισαγωγή σε κατάσταση ύπνου των στοιχείων του δικτύου που δεν είναι ενεργά, όπως παρουσιάζεται στον Πίνακα 1.5 με διαφορετικά επίπεδα μείωσης της ισχύος.

2) **Βελτιστοποίηση μεταξύ των κέντρων δεδομένων:** Πέρα από τις μεθόδους εξοικονόμησης ενέργειας εντός των κέντρων δεδομένων, έχουν προταθεί πολλές στρατηγικές για τη μείωση της κατανάλωσης ενέργειας μέσω της βέλτιστης κατανομής των δικτυακών φόρτων εργασίας ανάμεσα σε διαφορετικά, κατανεμημένα κέντρα δεδομένων. Αυτές οι προτάσεις αξιοποιούν τις διαφορές στα χαρακτηριστικά των κέντρων δεδομένων που βρίσκονται σε διάφορες γεωγραφικές περιοχές.

Εκτός από το υψηλό κόστος ηλεκτρικής ενέργειας, η κατανάλωση ενέργειας συμβάλλει σημαντικά και στις εκπομπές διοξειδίου του άνθρακα, προκαλώντας περιβαλλοντικές ζημιές. Συνήθως, η ηλεκτρική ενέργεια μιας περιοχής παράγεται από διάφορες γεννήτριες που χρησιμοποιούν διαφορετικά καύσιμα, όπως άνθρακα, πετρέλαιο, φυσικό αέριο, πυρηνική ενέργεια, αιολική και ηλιακή ενέργεια. Η παραγωγή ενέργειας από ορυκτά καύσιμα εκπέμπει πολύ περισσότερο διοξείδιο του άνθρακα σε σύγκριση με την ενέργεια που παράγεται από ανανεώσιμες πηγές. Για τη μείωση της συνολικής εκπομπής διοξειδίου του άνθρακα, έχουν αναπτυχθεί στρατηγικές βέλτιστης κατανομής των εργασιών σε «πράσινα» κέντρα δεδομένων.

1.3.7 Συντήρηση δικτύου

Αφού το δίκτυο του κέντρου δεδομένων εγκατασταθεί και τεθεί σε λειτουργία, είναι απαραίτητες συνεχείς εργασίες συντήρησης για την παρακολούθηση της κατάστασης του δικτύου, τον εντοπισμό προβλημάτων και την εκτέλεση απαραίτητων επισκευών. Οι αλλαγές ή ενημερώσεις στη διαμόρφωση είναι συνηθισμένες διαδικασίες στα κέντρα δεδομένων. Ωστόσο, αυτές οι διαδικασίες αποτελούν μία από τις κύριες αιτίες αστάθειας στο δίκτυο, όπως φαίνεται στον Πίνακα 1.6, και μπορεί να προκαλέσουν διακοπές και μειώσεις στην απόδοση.

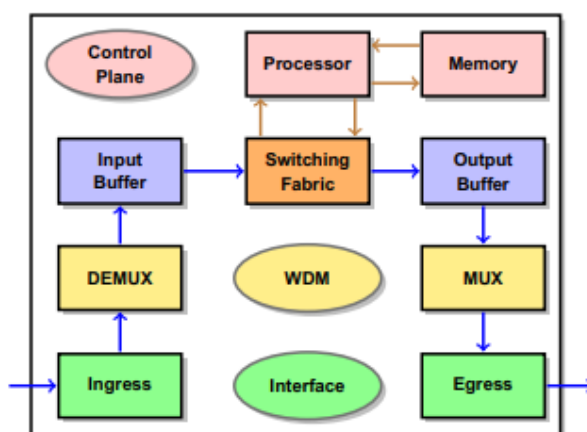
Category	Reason	Percentage
Software 21%	Link layer loop	19%
	Imbalance triggered overload	2%
Hardware 18%	FCS (frame checksum) error	13%
	Unstable power	5%
Configuration 38%	Errors on multiple switches	32%
	Errors on one switch	6%
	Switch stops forwarding	9%
Unknown 23%	Imbalance triggered overload	7%
	Lost configuration	5%
	High CPU utilization	2%

Πίνακας 1.6: Ταξινόμηση και στατιστικά στοιχεία των αστοχιών σε DCNs[1]

Κεφάλαιο 2 : Οπτική μεταγωγή σε κέντρα δεδομένων

2.1 Γενική αρχιτεκτονική οπτικού μεταγωγέα

Ένας τυπικός οπτικός switch περιλαμβάνει εξαρτήματα όπως φαίνεται στην εικόνα του Σχήματος 2.1. Οι εισοδοί και οι έξοδοι βασίζονται συνήθως σε πολυπλεξία διαίρεσης μήκους κύματος (WDM), επιτρέποντας τη μεταφορά πολλαπλών μηκών κύματος μέσω μιας ενιαίας ζεύξης. Κάθε μήκος κύματος μεταφέρει μια διαφορετική ροή δεδομένων. Ο αριθμός των μηκών κύματος, σύμφωνα με την τρέχουσα τεχνολογία, κυμαίνεται συνήθως γύρω στα 80, με απόσταση 50 GHz μεταξύ τους, και κάθε μήκος κύματος μπορεί να υποστηρίξει ταχύτητες δεδομένων μέχρι 40 Gbps .



Σχήμα 2.1 Εξαρτήματα ενός τυπικού οπτικού switch[2]

Σε αντίθεση με την αρχιτεκτονική μεταγωγής πακέτων, ο switch αποτελείται από πολυπλέκτες που διαχωρίζουν τα επιμέρους μήκη κύματος και αποπολυπλέκτες που συνδυάζουν πολλαπλά μήκη κύματος για τη μετάδοση μέσω μιας εξερχόμενης ίνας. Μια διάταξη μεταγωγής χρησιμοποιείται για να κατευθύνει τα δεδομένα από ένα εισερχόμενο μήκος κύματος στο αντίστοιχο μήκος κύματος εξόδου σε μια έξοδο. Απομονωτές μπορούν να χρησιμοποιηθούν για την προσωρινή αποθήκευση των πακέτων πριν αυτά προωθηθούν στον ιστό μεταγωγής. Σε έναν καθαρά οπτικό switch, το πακέτο μεταφέρεται αποκλειστικά σε οπτική μορφή. Ο ιστός μεταγωγής, σε αυτή την περίπτωση, είναι πλήρως οπτικός και δεν χρειάζεται να αναλύει την επικεφαλίδα ή άλλες πληροφορίες του πακέτου. Η μεταγωγή πραγματοποιείται με βάση τη σύνδεση εισόδου και το μήκος κύματος, κάτι που ονομάζεται αμιγώς οπτική μεταγωγή. Έχουν επίσης αναπτυχθεί αρχιτεκτονικές διακοπών που συνδυάζουν τεχνολογίες μεταγωγής πακέτων και οπτικής μεταγωγής, γνωστές ως υβριδικές αρχιτεκτονικές. Ορισμένες από αυτές

ενσωματώνουν έναν υβριδικό switch που συνδυάζει τόσο τα χαρακτηριστικά της μεταγωγής πακέτων όσο και της οπτικής μεταγωγής σε μια ενιαία συσκευή. Άλλες υλοποιούν μια διασύνδεση ανάμεσα σε μεταγωγείς πακέτων και οπτικούς μεταγωγείς. Στην επόμενη ενότητα θα εξεταστούν πιο αναλυτικά τα διάφορα συστατικά στοιχεία της μεταγωγής.

2.2.1 Επιλογές εξαρτημάτων Switch

Ένας οπτικός (ή υβριδικός) switch κατασκευάζεται συνδυάζοντας διάφορα εξαρτήματα. Κάθε ένα από αυτά τα εξαρτήματα μπορεί να υλοποιηθεί με διάφορες διαθέσιμες επιλογές. Αυτή η ενότητα παρέχει μια σύντομη παρουσίαση ορισμένων από τις σημαντικότερες επιλογές στοιχείων μεταγωγής που χρησιμοποιούνται στις προτεινόμενες αρχιτεκτονικές οπτικών ή υβριδικών διακοπών για δίκτυα δεδομένων (DCN).

2.2.2 Διεπαφή εισόδου-εξόδου

Το στοιχείο εισόδου-εξόδου (IO) λειτουργεί ως διεπαφή μεταξύ των εσωτερικών και εξωτερικών μερών ενός switch. Το πιο διαδεδομένο εξάρτημα που χρησιμοποιείται για αυτή τη λειτουργία είναι ο πομποδέκτης. Ο πομποδέκτης έχει ως κύριο ρόλο τη μετατροπή σημάτων από έναν τομέα σε έναν άλλο και αποτελείται από ένα ζεύγος πομπού και δέκτη που συνεργάζονται. Ο πομπός μετατρέπει τα σήματα από μια εσωτερική παράλληλη μορφή σε μια σειριακή ροή σημάτων, κατάλληλη για μετάδοση στο δίκτυο. Αντίθετα, ο δέκτης μετατρέπει τα σήματα που λαμβάνονται από το δίκτυο σε εσωτερική παράλληλη μορφή. Για παράδειγμα, ένας οπτικός δέκτης μετατρέπει τα οπτικά σήματα σε ηλεκτρονική μορφή που μπορεί να επεξεργαστεί ένας switch δικτύου. Στην αντίθετη κατεύθυνση, ένας πομπός μετατρέπει την ηλεκτρονική μορφή δεδομένων σε οπτικά σήματα που μεταδίδονται μέσω καλωδίου. Στη συζήτηση που ακολουθεί, θα επικεντρωθούμε αποκλειστικά σε οπτικούς πομποδέκτες, οι οποίοι χρησιμοποιούνται ευρέως τόσο σε συστήματα μεταγωγής πακέτων όσο και σε υβριδικά δίκτυα δεδομένων (όπως το LION). Ένας τυπικός πομπός αποτελείται από μια πηγή φωτός και έναν διαμορφωτή. Συνήθως, χρησιμοποιείται ένα λέιζερ ή μια δίοδος εκπομπής φωτός (LED) ως πηγή. Το φως που παράγεται από αυτή την πηγή διαμορφώνεται με βάση τα δεδομένα, μετατρέποντάς τα σε σήματα κατάλληλα για μετάδοση. Αντίστοιχα, ένας απλός δέκτης περιλαμβάνει έναν φωτοανιχνευτή (PD) και έναν ενισχυτή. Ο φωτοανιχνευτής ανιχνεύει την ύπαρξη ή απουσία φωτός και το μετατρέπει σε αντίστοιχα δεδομένα. Σε περιπτώσεις που τα σήματα είναι αδύναμα, ένας ενισχυτής χρησιμοποιείται για να τα ενισχύσει, έτσι ώστε όταν το σήμα ανιχνεύεται από τον φωτοανιχνευτή, τα αρχικά δεδομένα να ανακτώνται με λιγότερα σφάλματα. Αυτά τα εξαρτήματα καταναλώνουν ηλεκτρική ενέργεια για να λειτουργήσουν, και η κατανάλωση ισχύος τους εξαρτάται από τον ρυθμό μετάδοσης δεδομένων, το σχήμα κωδικοποίησης και τις επιλογές των επιμέρους στοιχείων. Συνήθως, ένας πομποδέκτης που λειτουργεί με υψηλότερο ρυθμό δεδομένων ή πιο σύνθετο σχήμα κωδικοποίησης καταναλώνει περισσότερη ενέργεια. Για τον ίδιο ρυθμό δεδομένων και το ίδιο σχήμα κωδικοποίησης, ένα λέιζερ καταναλώνει περισσότερη ισχύ από μια δίοδο εκπομπής φωτός (LED). Όλες οι αρχιτεκτονικές δικτύων δεδομένων (DCN) χρησιμοποιούν οπτικούς πομπούς, όπως δίοδοι εκπομπής LED και φωτοανιχνευτές (PD), για να υποστηρίξουν υψηλότερους ρυθμούς δεδομένων.

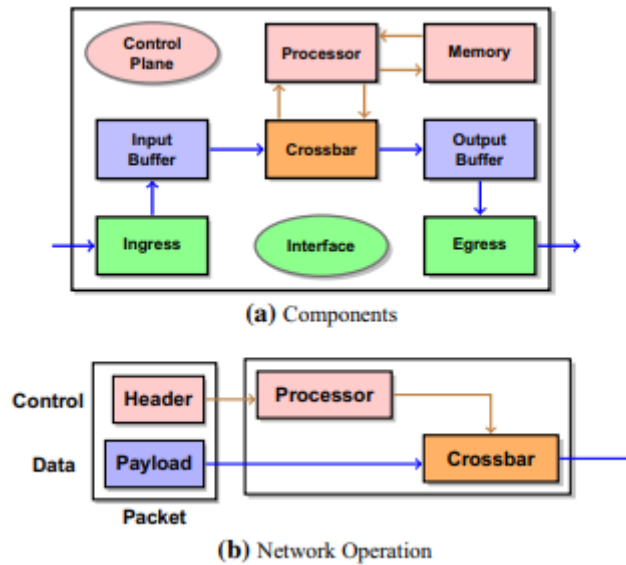
Το κόστος και η κατανάλωση ενέργειας ενός πομποδέκτη επηρεάζονται από τη δυνατότητα συντονισμού, την υποστήριξη μήκους κύματος και την απόσταση. Ένας πομποδέκτης με δυνατότητα συντονισμού απαιτεί περισσότερη ενέργεια σε σύγκριση με έναν πομποδέκτη σταθερού μήκους κύματος. Επίσης, ένας πομποδέκτης που λειτουργεί στα 1550 nm είναι κατάλληλος για επικοινωνίες μεγάλων αποστάσεων και συνήθως έχει υψηλότερο κόστος από έναν πομποδέκτη που λειτουργεί στα 850 nm. Ομοίως, ένας πομποδέκτης που συντονίζεται γρηγορότερα είναι ακριβότερος από έναν που συντονίζεται πιο αργά.

Όταν τα στοιχεία του switch λειτουργούν εξ ολοκλήρου στον οπτικό τομέα, δεν απαιτείται οπτικοηλεκτρική μετατροπή. Σε αυτή την περίπτωση, ένας απλός οπτικός σύνδεσμος αρκεί για τη μετάδοση των σημάτων χωρίς την ανάγκη πομποδέκτη. Όταν είναι απαραίτητη η μετατροπή μήκους κύματος, εξαρτήματα όπως ο μετατροπέας συντονιζόμενου μήκους κύματος (TWC) εκτελούν αυτή τη μετατροπή οπτικά, χωρίς να χρειάζεται η μετατροπή των σημάτων σε ηλεκτρονικό πεδίο. Αυτό το στοιχείο χρησιμοποιείται ευρέως σε οπτικές και υβριδικές αρχιτεκτονικές δικτύων δεδομένων (DCN).

Συνοπτικά, παρουσιάστηκαν τέσσερις βασικές επιλογές διεπαφών εισόδου-εξόδου: οπτικός πομποδέκτης σταθερού μήκους κύματος, συντονιζόμενος πομποδέκτης, σύνδεσμος οπτικών ινών και TWC. Στην επόμενη υποενότητα θα εξεταστούν πιθανές επιλογές υλοποίησης για το δίκτυο μεταγωγής (SwF).

2.2.3 Δίκτυο μεταγωγής (Switch Fabric)

Το SwF (Switch Fabric) παρέχει την υποδομή για τη διασύνδεση εντός ενός switch, επιτρέποντας τη μεταφορά των πακέτων από τη θύρα εισόδου στη θύρα εξόδου προορισμού. Οι συνήθεις τοπολογίες που χρησιμοποιούνται για την σύνδεση ενός SwF με τις θύρες εισόδου και εξόδου είναι ο δίαυλος και το σημείο προς σημείο. Η τοπολογία δίαυλου δεν κλιμακώνεται καλά όσον αφορά την απόδοση, γι' αυτό για υψηλότερη απόδοση χρησιμοποιούνται συνδέσεις σημείο προς σημείο.



Σχήμα 2.2 Εξαρτήματα και λειτουργία ενός τυπικού μεταγωγέα πακέτων[2]

Μια διασταυρούμενη υποδομή (όπως φαίνεται στο Σχήμα 2.2) επιτρέπει τη μεταφορά πολλών πακέτων ταυτόχρονα μέσω του SwF. Κατά τη διάρκεια της διαδικασίας μεταφοράς, διατηρείται ένα σύνολο γραμμών δεδομένων σε υψηλές και χαμηλές καταστάσεις τάσης για την αναπαράσταση των μονάδων και των μηδενικών. Υπάρχει σημαντική κατανάλωση ισχύος για τη διατήρηση αυτών των γραμμών στις αντίστοιχες καταστάσεις τους, σύμφωνα με την είσοδο των δεδομένων του πακέτου. Αυτό είναι παρόμοιο με τις πολλές γραμμές που συνδέονται με μια μονάδα μνήμης. Καθώς αυξάνεται ο αριθμός των γραμμών, η κατανάλωση ισχύος αυξάνεται αναλογικά. Ομοίως, καθώς ο αριθμός των θυρών αυξάνεται, η κατανάλωση ισχύος των διασταυρούμενων συνδέσεων αυξάνεται με τετραγωνική αναλογία.

Όταν πολλά πακέτα επιλέγουν ταυτόχρονα την ίδια θύρα εξόδου, προκύπτει ανταγωνισμός. Για την αντιμετώπιση αυτού του ανταγωνισμού, ενσωματώνονται μικρές απομονωτικές μνήμες στην υποδομή διασύνδεσης. Αυτές οι διασταυρούμενες υποδομές (SwF) καταναλώνουν περισσότερη ενέργεια σε σύγκριση με εκείνες που δεν διαθέτουν ρυθμιστές.

Ένα SwF στο οπτικό τομέα μπορεί να υλοποιηθεί με συνδυασμό των παρακάτω οπτικών στοιχείων:

1) Ο **διαχωριστής/συνδυαστής (SC)** λειτουργεί με τρόπο παρόμοιο με τη διασύνδεση διαύλου στον οπτικό τομέα. Κάθε σήμα που εισέρχεται στον SC διανέμεται σε όλες τις θύρες εξόδου.

2) Το **μικροηλεκτρομηχανικό σύστημα (MEMS)** χρησιμοποιεί ηλεκτρομηχανική ευθυγράμμιση κατόπτρων για τη σύζευξη οπτικών σημάτων στο χωρικό πεδίο, παρόμοια με τη διασταυρούμενη σύνδεση στον ηλεκτρονικό τομέα. Παρ' όλα αυτά, ο χρόνος που απαιτείται για την αναδιαμόρφωση αυτών των διακοπών είναι της τάξης των χιλιοστών του δευτερολέπτου.

3) Οι **switch δρομολογητών με διάταξη κυματοδηγών (AWGR)** λειτουργούν ταυτόχρονα τόσο στο χωρικό πεδίο όσο και στο πεδίο του μήκους κύματος. Τα μήκη κύματος $\{\lambda_1, \dots, \lambda_m\}$ από μια θύρα εισόδου κατευθύνονται προς τις αντίστοιχες θύρες εξόδου $\{p_1, \dots, p_m\}$. Μπορεί να σχεδιαστεί ώστε ένα συγκεκριμένο μήκος κύματος (π.χ. λ_1) από μια άλλη θύρα εισόδου να δρομολογείται σε διαφορετική θύρα εξόδου ($p_i = p_1$). Έτσι, επιτυγχάνεται ταυτόχρονη δρομολόγηση σημάτων από την είσοδο στις θύρες εξόδου τόσο στο χωρικό όσο και στο πεδίο του μήκους κύματος.

4) Η **οπτική επίλυση διαφωνιών με ηλεκτρονικούς απομονωτές (OCE)** χρησιμοποιεί μικρούς ηλεκτρονικούς απομονωτές για την επίλυση του ανταγωνισμού σε συνδυασμό με ένα AWGR. Αυτή η προσέγγιση μοιάζει με την ηλεκτρονική στρατηγική SwF.

5) Ο **επιλεκτικός switch μήκους κύματος (WSS)** λειτουργεί τόσο στο πεδίο του μήκους κύματος όσο και στο χώρο. Μπορεί να ρυθμιστεί ώστε να επιλέγει ένα ή περισσότερα μήκη κύματος από την είσοδο και να τα κατευθύνει προς μια θύρα απόρριψης. Τα υπόλοιπα μήκη κύματος συνεχίζουν να περνούν μέσω της θύρας express. Αντίστοιχα, μια θύρα προσθήκης συνδυάζει τα επιλεγμένα μήκη κύματος από μια θύρα εισόδου προσθήκης με αυτά της θύρας express. Έτσι, οι θύρες προσθήκης και απόρριψης λειτουργούν συμπληρωματικά, επιτρέποντας τη δρομολόγηση συγκεκριμένων μηκών κύματος όπου είναι αναγκαίο.

Από τις οπτικές επιλογές SwF, τα SC και AWGR δεν απαιτούν κατανάλωση ηλεκτρικής ενέργειας. Η μονάδα MEMS χρειάζεται κυρίως ηλεκτρική ενέργεια για την πραγματοποίηση αναδιαμορφώσεων. Η OCE με ηλεκτρονικούς απομονωτές χρησιμοποιεί τον απομονωτή σε περίπτωση ανταγωνισμού. Ωστόσο, αυτό το κύκλωμα απαιτεί έναν πομποδέκτη για τη μετατροπή των σημάτων από το οπτικό στο ηλεκτρονικό πεδίο, καθιστώντας το πιο περίπλοκο και με υψηλότερη κατανάλωση ενέργειας σε σύγκριση με άλλες οπτικές επιλογές SwF. Ένα crossbar χρησιμοποιείται σε δίκτυα κέντρων δεδομένων με μεταγωγή πακέτων, ενώ τα υπόλοιπα οπτικά στοιχεία SwF μπορούν να εφαρμοστούν τόσο σε οπτικές όσο και σε υβριδικές αρχιτεκτονικές DCN.

2.2.4 Port buffers

Οι απομονωτές θύρας εισόδου χρησιμοποιούνται για να συγκρατούν τα πακέτα μέχρι να είναι έτοιμο το SwF να τα προωθήσει στη θύρα εξόδου. Οι απομονωτές θύρας εξόδου είναι συνήθως μικροί και διατηρούν το πακέτο μέχρι η εξερχόμενη ζεύξη να είναι έτοιμη για τη μετάδοσή του. Μια κοινή πρακτική είναι η χρήση τόσο των απομονωτών θύρας εισόδου όσο και εξόδου. Στην ιδανική περίπτωση, το SwF είναι έτοιμο να μεταφέρει ένα πακέτο αμέσως μόλις αυτό φτάσει, γεγονός που συμβαίνει μόνο όταν ο αντίστοιχος απομονωτής θύρας εισόδου είναι άδειος. Διαισθητικά, όταν τα εισερχόμενα πακέτα φτάνουν με τέτοιο χρονισμό ώστε κάθε πακέτο να συναντά πάντα έναν άδειο απομονωτή, η απαιτούμενη χωρητικότητα του αντίστοιχου απομονωτή μπορεί να μηδενιστεί. Αυτή η στρατηγική εφαρμόζεται στην αρχιτεκτονική Fastpass. Ωστόσο, λόγω της έλλειψης συγχρονισμού, αυτή η αρχιτεκτονική απαιτεί τη χρήση μικρών απομονωτών.

Στον οπτικό τομέα, οι γραμμές καθυστέρησης (ή οπτικές ίνες συγκεκριμένου μήκους) χρησιμοποιούνται για την εισαγωγή μιας συγκεκριμένης καθυστέρησης στη διάδοση του σήματος. Για παράδειγμα, μια οπτική ίνα 40 μέτρων μπορεί να καθυστερήσει ένα πακέτο κατά 200 ns. Μέσα σε αυτό το χρονικό διάστημα, η διαδικασία αναζήτησης για τον καθορισμό της

διαδρομής του εισερχόμενου πακέτου πρέπει να ολοκληρωθεί, ώστε το πακέτο να μπορεί να εξυπηρετηθεί άμεσα από το SwF.

Υπάρχουν τέσσερις επιλογές για τους απομονωτές θύρας:

- 1) Μεγάλες ηλεκτρονικές αποθήκες θύρας
- 2) Μικρές αποθήκες θύρας
- 3) Μηδενικοί απομονωτές (καθόλου απομονωτές θύρας)
- 4) Οπτικές γραμμές καθυστέρησης

Από αυτές τις επιλογές, οι τρεις πρώτες αφορούν ειδικά τη μεταγωγή δεδομένων σε δίκτυα πακέτων, ενώ οι οπτικές γραμμές καθυστέρησης χρησιμοποιούνται σε οπτικές και υβριδικές αρχιτεκτονικές DCN. Η επιλογή αυτή επηρεάζει την απόδοση του δικτύου όσον αφορά την καθυστέρηση των πακέτων, αν και η επίδραση στην απόδοση δεν είναι ιδιαίτερα σημαντική. Από την άποψη της κατανάλωσης ενέργειας, οι επιλογές (1) και (2) καταναλώνουν σημαντική ποσότητα ισχύος, ενώ οι άλλες δύο επιλογές δεν καταναλώνουν ενέργεια.

2.2.5.1 Αναζήτηση επικεφαλίδας(Header Lookup)

Η συνάρτηση lookup υπολογίζει τη θύρα εξόδου για ένα εισερχόμενο πακέτο, βασισμένη είτε στα δεδομένα της επικεφαλίδας του πακέτου είτε σε μεταδεδομένα του πακέτου, όπως η εισερχόμενη θύρα ή/και το μήκος κύματος εισόδου. Συνήθως χρησιμοποιούμενοι μηχανισμοί περιλαμβάνουν:

- 1) Αναζήτηση πακέτου
- 2) Αναζήτηση ετικέτας
- 3) Αναζήτηση OpenFlow
- 4) Αναζήτηση μήκους κύματος

Κάθε μηχανισμός αναζήτησης έχει διαφορετικές απαιτήσεις σε μνήμη και επεξεργαστικούς πόρους, οι οποίες θα αναλυθούν στη συνέχεια.

2.2.5.2 Αναζήτηση πακέτων

Η αναζήτηση πακέτων είναι ο πιο διαδεδομένος μηχανισμός αναζήτησης και περιλαμβάνει μία ή περισσότερες διαδικασίες αναζήτησης. Σε ένα τυπικό σενάριο δρομολόγησης, η αναζήτηση γίνεται στην επικεφαλίδα IP για τον καθορισμό του επόμενου άλματος. Η διαδικασία περιλαμβάνει τα εξής βήματα:

1) Αποκαθήλωση του Ωφέλιμου Φορτίου και Εξαγωγή Κατώτερων Επιπέδων: Αρχικά, αφαιρείται το ωφέλιμο φορτίο και εξάγονται οι επικεφαλίδες κατώτερων επιπέδων. Για παράδειγμα, σε μια περίπτωση IP πάνω από Ethernet, οι επικεφαλίδες Ethernet αφαιρούνται από το εισερχόμενο πακέτο.

2) Έλεγχος και Επικύρωση Αθροίσματος Ελέγχου: Το αθροιστικό έλεγχο εξάγεται και επικυρώνεται για να διασφαλιστεί η ακεραιότητα του πακέτου. Αφού ολοκληρωθεί αυτή η διαδικασία, το ωφέλιμο φορτίο του πακέτου IP είναι έτοιμο για επεξεργασία.

3) Επεξεργασία Επικεφαλίδων IP: Το ωφέλιμο φορτίο του πακέτου IP περιλαμβάνει τις επικεφαλίδες IP που υποδεικνύουν τον προορισμό του πακέτου.

4) Αναζήτηση στον Πίνακα Δρομολόγησης: Οι πληροφορίες από τις επικεφαλίδες IP χρησιμοποιούνται για την αναζήτηση του μεγαλύτερου προθέματος που ταιριάζει από το σύνολο των καταχωρίσεων του πίνακα δρομολόγησης.

Αυτή η αναζήτηση απαιτεί τη χρήση ενός πρωτοκόλλου δρομολόγησης, όπως το RIP ή το OSPF, για την ανταλλαγή ενημερώσεων δρομολόγησης μέσω του δικτύου. Με βάση αυτές τις ενημερώσεις, δημιουργείται και διατηρείται ο πίνακας δρομολόγησης IP. Η διαδικασία δρομολόγησης απαιτεί σημαντική ποσότητα μνήμης και επεξεργαστικών πόρων για τη συντήρηση του πίνακα. Ο πίνακας δρομολόγησης είναι ένας πίνακας βασισμένος σε εύρος τιμών και επομένως απαιτεί αντιστοίχιση με το μεγαλύτερο πρόθεμα για κάθε πακέτο. Αυτή η αντιστοίχιση με το μεγαλύτερο πρόθεμα μπορεί να προκαλέσει σημαντική καθυστέρηση. Για να επιταχυνθούν οι αναζητήσεις, χρησιμοποιούνται μονάδες διευθυνσιοδοτούμενης μνήμης τριτογενούς περιεχομένου (TCAM), οι οποίες μειώνουν την καθυστέρηση της αναζήτησης, αλλά καταναλώνουν περισσότερη ενέργεια. Συνοψίζοντας, η αναζήτηση πακέτων απαιτεί σημαντική ποσότητα μνήμης και επεξεργασίας, και επηρεάζει την απόδοση του δικτύου, κυρίως όσον αφορά την καθυστέρηση.

2.2.5.3 Αναζήτηση OpenFlow

Η αναζήτηση OpenFlow έχει κερδίσει ευρεία αποδοχή λόγω της ευελιξίας που προσφέρει. Χρησιμοποιεί έναν κεντρικό ελεγκτή που υπολογίζει τη διαδρομή για μια δεδομένη ροή και εισάγει την αντίστοιχη καταχώρηση ροής στους πίνακες προώθησης. Οι μεταγωγείς στη συνέχεια αναζητούν σε αυτούς τους πίνακες για να καθορίσουν την κατεύθυνση του πακέτου.

Η διαδικασία αναζήτησης OpenFlow περιλαμβάνει τα εξής βήματα:

1) Εξαγωγή Πεδίων από το Εισερχόμενο Πακέτο: Τα πεδία του πακέτου, όπως η διεύθυνση IP προορισμού, η ετικέτα Ethernet VLAN και η διεύθυνση IP πηγής, εξάγονται από το εισερχόμενο πακέτο.

2) Χρήση Πεδίων ως Κλειδιά: Αυτά τα πεδία χρησιμοποιούνται ως κλειδιά για την αναζήτηση σε έναν ή περισσότερους πίνακες, ακολουθώντας μια προκαθορισμένη σειρά αναζήτησης.

3) Ανάλυση Αποτελέσματος Αναζήτησης: Το αποτέλεσμα της αναζήτησης μπορεί να είναι είτε μια καταχώρηση σε άλλο πίνακα είτε ένας συνδυασμός ενεργειών και καταχωρήσεων που καθορίζουν την επόμενη ενέργεια για το πακέτο.

Αυτή η διαδικασία επιτρέπει την ευέλικτη και λεπτομερή διαχείριση της ροής δεδομένων, εξασφαλίζοντας ότι κάθε πακέτο κατευθύνεται σωστά βάσει των καθορισμένων κανόνων. Επιπλέον, αυτός ο μηχανισμός αναζήτησης περιλαμβάνει έναν μεταβλητό αριθμό πινάκων και είναι ελαφρώς πιο πολύπλοκος από μια απλή αναζήτηση ετικετών. Ωστόσο, το κύριο πλεονέκτημα είναι ότι, στην αναζήτηση OpenFlow, η επιλογή των πεδίων ανήκει στον διαχειριστή του δικτύου, παρέχοντας μεγαλύτερη ευελιξία και έλεγχο στον τρόπο διαχείρισης της ροής δεδομένων. Αντίθετα, στην αναζήτηση πακέτων, η επιλογή των πεδίων καθορίζεται από τον προμηθευτή του εξοπλισμού, περιορίζοντας την ευελιξία του δικτύου.

2.2.5.4 Αναζήτηση οπτικής μεταγωγής πακέτων (OPS)

Αυτός ο μηχανισμός περιλαμβάνει γρήγορη και χρονικά περιορισμένη εξαγωγή ετικέτας, όπως χρησιμοποιείται στο σύστημα LIONS. Αυτή η διαδικασία είναι παρόμοια με την αναζήτηση ετικετών, αλλά με την αυστηρή απαίτηση να ολοκληρωθεί εντός συγκεκριμένου χρονικού περιθωρίου. Ο switch LIONS συνήθως έχει αυστηρή προθεσμία 190 ns για να ολοκληρώσει την αναζήτηση και την αναδιαμόρφωση της ετικέτας. Εάν η προθεσμία δεν τηρηθεί, το πακέτο χάνεται και πρέπει να επαναμεταδοθεί στον μεταγωγέα LIONS. Αυτό απαιτεί υψηλή ταχύτητα επεξεργασίας και μνήμη για την εξαγωγή και επεξεργασία της ετικέτας.

2.2.5.5 Αναζήτηση μήκους κύματος

Σε αυτόν τον μηχανισμό αναζήτησης, το μήκος κύματος χρησιμοποιείται ως δείκτης διαδρομής. Έτσι, όταν ένα οπτικό σήμα από μια εισερχόμενη θύρα p_1 φτάνει σε ένα συγκεκριμένο μήκος κύματος (π.χ. λ_1), προωθείται στην αντίστοιχη θύρα εξόδου p_i . Αυτός ο μηχανισμός δεν απαιτεί ρητή εξαγωγή ετικέτας και στηρίζεται στη δρομολόγηση μήκους κύματος στον οπτικό τομέα. Ο κύκλωμα μήκους κύματος είναι προκαθορισμένο, επομένως δεν απαιτούνται επιπλέον πόροι μνήμης ή επεξεργασίας για αυτήν την αναζήτηση.

Η συνοπτική παρουσίαση των μηχανισμών αναζήτησης είναι διαθέσιμη στον Πίνακα 2.1. Από τους διάφορους μηχανισμούς αναζήτησης, οι δύο τελευταίοι, δηλαδή η αναζήτηση βάσει μήκους κύματος και η αναζήτηση οπτικού πακέτου, χρησιμοποιούνται σε οπτικές και υβριδικές αρχιτεκτονικές δικτύων κέντρων δεδομένων. Αντίθετα, άλλοι μηχανισμοί αναζήτησης είναι ειδικά σχεδιασμένοι για μεταγωγείς πακέτων.

Lookup	Memory	Processing	Latency	Set-up
Packet	Medium	High	High	Not required
Label	High	Medium	Medium	Required
OpenFlow	High ^a	Medium ^a	Medium ^a	Required
OPS	High	High	Low	Not required
Wavelength	Low	Low	Low	Required

^a Open-Flow resource requirements are variable and when not managed can exceed all other lookup mechanisms

Πίνακας 2.1: Σύγκριση των μηχανισμών αναζήτησης όσον αφορά τη μνήμη τους, απαιτήσεις σε πόρους επεξεργασίας, την καθυστέρηση και κατά πόσον η αρχική ρύθμιση απαιτείται[2]

2.2.5.6 Σύνοψη εξαρτημάτων Switch

Μια σύνοψη όλων των επιλογών εξαρτημάτων που έχουν παρουσιαστεί μέχρι στιγμής παρουσιάζεται στον πίνακα 2.2. Παρουσιάζονται η κατανάλωση ισχύος, η μέγιστη απόδοση, η καθυστέρηση στο επίπεδο δεδομένων και η καθυστέρηση στο επίπεδο ελέγχου.

Στον πίνακα, η παράμετρος O_{max} είναι η μέγιστη οπτική εύρος ζώνης που υποστηρίζεται από ένα σκέλος ίνας, μ είναι ο ρυθμός δεδομένων της σύνδεσης(link data rate), $|I|$ είναι ο αριθμός των θυρών εισόδου, $|O|$ είναι ο αριθμός των θυρών εξόδου και $|B|$ είναι το πλάτος διαύλου ενός διασταυρομένου (crossbar) SwF.

Attribute→		Power	Throughput	Delay	
Component ↓	Choice			Data	Control
Interface	Receiver	$\propto \mu$	$= \mu$	$\propto \mu L$	0
	Transmitter	$\propto \mu$	$= \mu$	$\propto \mu L$	0
	Connector	0	O_{max}	0	0
	TWC and TL	C	O_{max}	0	10 ns
Port buffers	Zero	0	O_{max}	0	0
	Small	0.5 W	C_{max}	Variable	0
	Large	2 W	C_{max}	Variable	0
	Delay line	0	O_{max}	190 ns	0
SwF	Crossbar	$\propto I O B $	C_{max}	C	0
	AWGR	0	O_{max}	0	0
	MEMS	45 W	O_{max}	0	10 ms
	SC	0	O_{max}	0	0
	WSS	6–8 W	O_{max}	0	10 ms

Delay in data plane and control plane are separately mentioned. Here O_{max} is the maximum optical bandwidth of a fiber strand considering the encoding scheme, C_{max} is the crossbar bandwidth, C is a constant and μ is the corresponding data rate. Sets I , O and B correspond to the input ports, output ports and bus lines of a crossbar, respectively

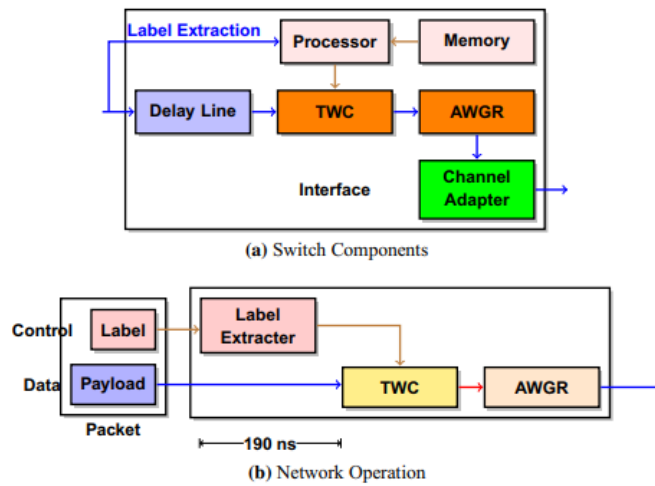
Πίνακας 2.2 : Κατανάλωση ενέργειας, μέγιστη απόδοση και καθυστέρηση διαφορετικών επιλογών εξαρτημάτων[2]

2.3.1 Οπτικές αρχιτεκτονικές DCN

Σε αυτή την ενότητα παρουσιάζονται αντιπροσωπευτικά παραδείγματα οπτικών και υβριδικών αρχιτεκτονικών DCN. Συγκεκριμένα, οι αρχιτεκτονικές αυτές επιλέγονται βάσει της ποικιλίας των χρησιμοποιούμενων στοιχείων. Επιπλέον, τα στοιχεία του υβριδικού switch στις αντίστοιχες αρχιτεκτονικές έχουν χρωματιστεί για να υπογραμμίζουν τους ρόλους τους. Τα εξαρτήματα διασύνδεσης εμφανίζονται με πράσινο χρώμα, τα εξαρτήματα του ρυθμιστικού διαύλου με μπλε, τα στοιχεία SwF με πορτοκαλί και τα στοιχεία του επιπέδου ελέγχου με καφέ χρώμα.

2.3.2 Αρχιτεκτονική LION

Το οπτικό δίκτυο διασύνδεσης φωτός (LION) αποτελεί μια υβριδική αρχιτεκτονική switch. Αυτή η αρχιτεκτονική ενσωματώνει έναν υβριδικό switch και χρησιμοποιεί το AWGR ως το στοιχείο SwF. Τα βασικά στοιχεία ενός LION απεικονίζονται στο σχήμα 2.3α. Επιπλέον, αυτή η αρχιτεκτονική χρησιμοποιεί ένα TWC ως το στοιχείο διασύνδεσης, το οποίο τοποθετείται μπροστά από τον SwF. Η διαδικασία αναζήτησης βασίζεται στην εξαγωγή επικεφαλίδων, κάτι που θυμίζει τον μηχανισμό που χρησιμοποιείται στους οπτικούς μεταγωγείς πακέτων.



Σχήμα 2.3 Στοιχεία και λειτουργία ενός LION switch[2]

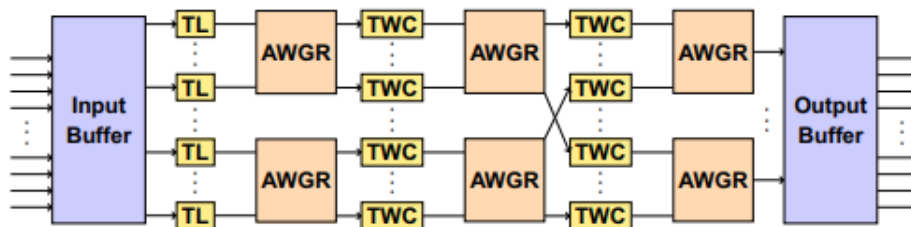
Η λειτουργία του δικτύου μοιάζει με εκείνη ενός οπτικού μεταγωγέα πακέτων, όπως φαίνεται στο Σχήμα 2.3β. Όταν παραλαμβάνεται ένα πακέτο, εκτελούνται τα εξής βήματα :

- 1) Το εισερχόμενο πακέτο διαχωρίζεται σε δύο διαδρομές. Η μία διαδρομή οδηγεί στην εξαγωγή της επικεφαλίδας, ενώ η άλλη κατευθύνεται προς μια γραμμή καθυστέρησης.
- 2) Η επικεφαλίδα εξάγεται από τον επεξεργαστή, ο οποίος υπολογίζει την κατάλληλη θύρα εξόδου. Παράλληλα, εντοπίζεται ένα διαθέσιμο μήκος κύματος και θύρα, μέσω των οποίων θα μπορέσει το πακέτο να περάσει στην αντίστοιχη εξερχόμενη θύρα.
- 3) Σε περίπτωση ανταγωνισμού, υπολογίζεται ένα ελεύθερο μήκος κύματος για τη δρομολόγηση του πακέτου προς τους ηλεκτρονικούς απομονωτές. Αυτός είναι ο μηχανισμός επίλυσης ανταγωνισμού που εφαρμόζεται. Σε έναν επόμενο κύκλο, το πακέτο επανεμφανίζεται από τον εσωτερικό απομονωτή πακέτων.
- 4) Ο TWC συντονίζεται στο ελεύθερο μήκος κύματος που έχει υπολογιστεί στα προηγούμενα βήματα.
- 5) Μετά τον συντονισμό, το πακέτο καθυστερεί στη δεύτερη διαδρομή μέσω της παθητικής γραμμής καθυστέρησης και παρουσιάζεται ως είσοδος στον συντονισμένο μετατροπέα μήκους κύματος.
- 6) Ο μετατροπέας αλλάζει το πακέτο στο ελεύθερο μήκος κύματος (κόκκινο). Το πακέτο εξόδου δρομολογείται στο οπτικό πεδίο μέσω του AWGR και καταλήγει στην εξερχόμενη θύρα.
- 7) Η εξερχόμενη θύρα είναι εξοπλισμένη με οπτοηλεκτρονικούς μετατροπείς, οι οποίοι μετατρέπουν την έξοδο πολλαπλών μηκών κύματος σε ηλεκτρονικό σήμα, το οποίο στη συνέχεια μεταφέρεται σε μία μόνο εξερχόμενη διεπαφή.

Αξίζει να σημειωθεί ότι ο χρονισμός παίζει καθοριστικό ρόλο στη λειτουργία αυτού του switch. Η γραμμή καθυστέρησης εισάγει ακριβώς την απαιτούμενη καθυστέρηση, η οποία αντιστοιχεί στον χρόνο που χρειάζεται για την εξαγωγή της επικεφαλίδας και τη διαμόρφωση του συντονιζόμενου μετατροπέα μήκους κύματος. Αυτό λειτουργεί εντός ενός χρονικού πλαισίου περίπου 190 ns.

2.3.3 Αρχιτεκτονική Petabit

Η αρχιτεκτονική Petabitswitch, χρησιμοποιεί έναν υβριδικό switch με πολλαπλά στάδια AWGRs για την υλοποίηση ενός κλιμακούμενου συστήματος, όπως φαίνεται στο Σχήμα 2.4. Αρκετά AWGR συνδυάζονται για να σχηματίσουν ένα ενιαίο στάδιο. Το σύστημα αυτό διαμορφώνεται ως ένα μη μπλοκαρισμένο δίκτυο Clos με m στάδια ($m = 1 + \log_2 N\gamma$), όπου N και γ είναι ο αριθμός των θυρών σε ένα AWGR και στον μεταγωγέα Petabit, αντίστοιχα. Προστίθεται ένα επιπλέον στάδιο για τη μονάδα εισόδου. Στη μονάδα εισόδου, ρυθμιζόμενα λέιζερ (TL) καθοδηγούν ένα ή περισσότερα πακέτα προς την επιθυμητή θύρα εξόδου. Σε άλλα στάδια, οι μετατροπείς συντονιζόμενου μήκους κύματος αναλαμβάνουν την καθοδήγηση των πακέτων. Οι ρυθμιστές εισόδου και εξόδου διαχειρίζονται την ένωση και τον διαχωρισμό πολλαπλών πακέτων που κατευθύνονται στην ίδια θύρα εξόδου, χρησιμοποιώντας ξεχωριστή εικονική ουρά εξόδου για κάθε θύρα προορισμού.

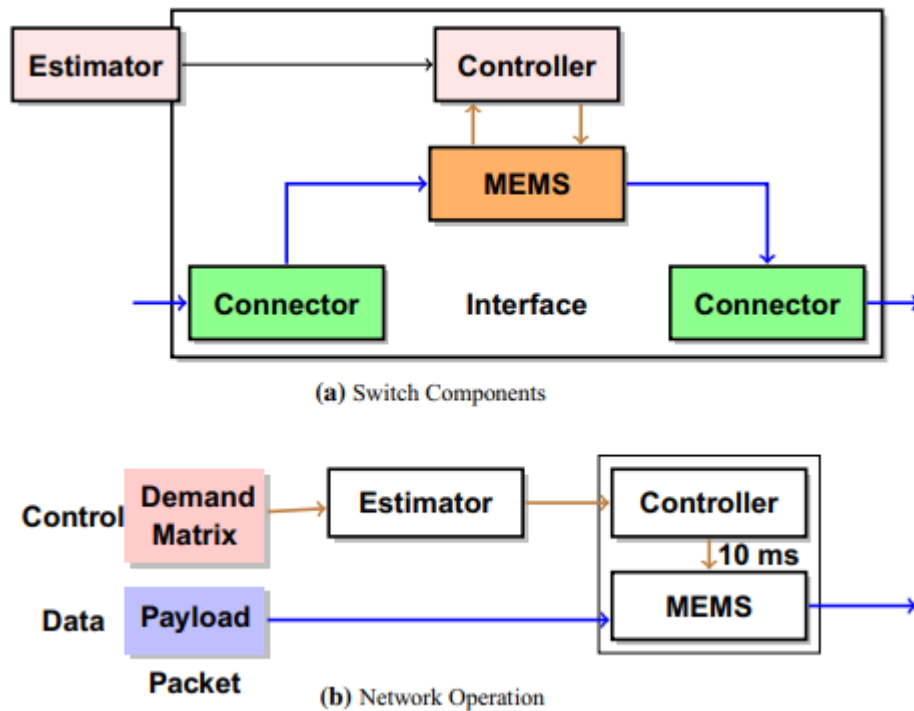


Σχήμα 2.4 Αρχιτεκτονική μεταγωγέα Petabit με $m + 1$ στάδια δρομολογητών AWGR[2]

Η λειτουργία του δικτύου του μεταγωγέα Petabit είναι παρόμοια με αυτή του LION. Στην αρχιτεκτονική LION, ο μεταγωγέας πρέπει να υπολογίζει το χρονοδιάγραμμα εντός αυστηρού χρονικού πλαισίου 190 ns για κάθε AWGR. Αντίθετα, στον μεταγωγέα Petabit, οι απομονωτές εισόδου και οι εικονικές ουρές εξόδου μπορούν να διατηρούν τα πακέτα μέχρι να βρεθεί και να υπολογιστεί μια διαδρομή χωρίς ανταγωνισμό. Αυτή η διαδρομή καθορίζεται μέσα από πολλαπλά στάδια. Επιπλέον, τα ρυθμιζόμενα στοιχεία πρέπει να ελέγχονται συγχρονισμένα από τον μεταγωγέα για να εξασφαλιστεί η ομαλή λειτουργία του συστήματος.

2.3.4 Αρχιτεκτονική Helios

Η αρχιτεκτονική Helios, η οποία βασίζεται στη μεταγωγή κυκλώματος, πρόκειται για μια υβριδική αρχιτεκτονική δικτύου που συνδυάζει μεταγωγείς πακέτων και οπτικούς μεταγωγείς. Ένα MEMS χρησιμοποιείται ως SwF στον οπτικό switch της. Δεν περιλαμβάνει μηχανισμό αναζήτησης και διαθέτει μόνο συνδέσμους. Μια εισερχόμενη θύρα είναι προκαθορισμένη να μεταβιβάζει όλα τα εισερχόμενα πακέτα σε μια συγκεκριμένη θύρα εξόδου. Αυτή η διαμόρφωση επανεξετάζεται περιοδικά, προσαρμόζοντας τη διανομή της κυκλοφορίας σύμφωνα με τις τρέχουσες ανάγκες.

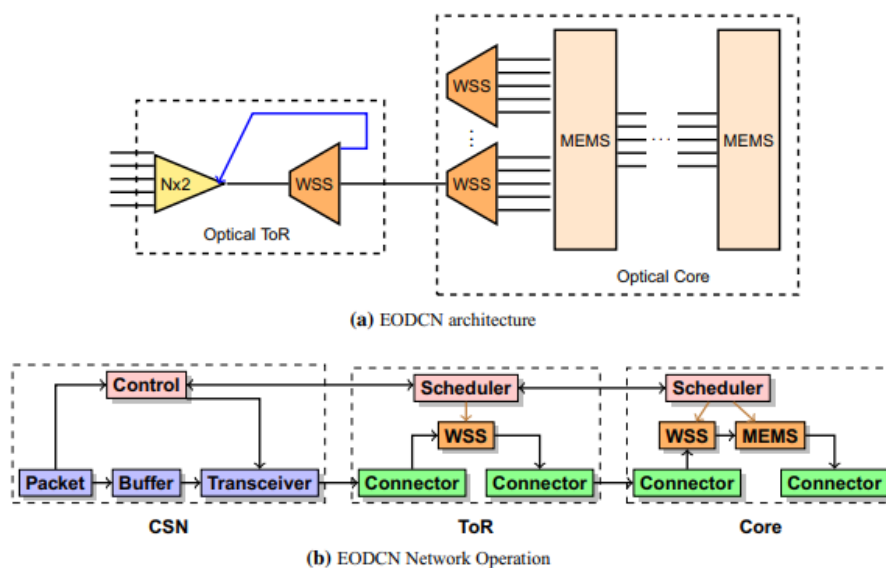


Σχήμα 2.5 Στοιχεία και λειτουργία μιας Helios αρχιτεκτονικής[2]

Τα στοιχεία και η λειτουργία της αρχιτεκτονικής Helios απεικονίζονται στο Σχήμα 2.5. Οι κυρίαρχες ροές ζήτησης εντοπίζονται από έναν εκτιμητή. Η αντιστοίχιση εισόδου και εξόδου αυτών των ροών πραγματοποιείται με τη βοήθεια ενός αλγορίθμου μέγιστου ταιριάσματος. Η προκύπτουσα διαμόρφωση προωθείται περιοδικά στο MEMS μέσω ενός ελεγκτή. Η αναδιαμόρφωση του στοιχείου MEMS απαιτεί περίπου 10 ms. Μετά την ολοκλήρωση αυτού του βήματος, οι κυρίαρχες ροές μπορούν να τροποποιηθούν πλήρως στον οπτικό τομέα. Οποιαδήποτε άλλη ροή τροποποιείται στο ηλεκτρονικό πεδίο χρησιμοποιεί ένα σύνολο μεταγωγέων πακέτων (packet switches).

2.3.5 Αρχιτεκτονική EODCN

Η αρχιτεκτονική του ελαστικού οπτικού DCN (EODCN), συνδυάζει δρομολόγηση μήκους κύματος, μεταγωγή κυκλώματος και μηχανισμούς μετάδοσης και επιλογής. Αποτελεί μια αρχιτεκτονική οπτικού δικτύου κέντρου δεδομένων που δεν χρησιμοποιεί στοιχεία μεταγωγέων πακέτων. Εισάγει έναν οπτικό κόμβο ToR που χρησιμοποιεί έναν οπτικό ζεύκτη $N \times 2$, όπως φαίνεται στο Σχήμα 2.6α. Η έξοδος αυτού του ζεύκτη τροφοδοτείται σε ένα WSS (Wavelength Selective Switch). Οι επιλογές μήκους κύματος από το WSS συνδέονται πίσω στο rack για ενδο-rack επικοινωνία. Αυτός ο οπτικός ToR συμβάλλει στη σημαντική μείωση της κατανάλωσης ενέργειας. Το μόνο ηλεκτρικά ενεργό στοιχείο σε αυτόν τον οπτικό μεταγωγέα ToR είναι το WSS, το οποίο καταναλώνει λιγότερο από 10 W, σε σύγκριση με τους εμπορικούς μεταγωγείς ToR που καταναλώνουν αρκετές εκατοντάδες watt.



Σχήμα 2.6 Αρχιτεκτονική και λειτουργία του EODCN[2]

Ένα ελκυστικό χαρακτηριστικό αυτού του μεταγωγέα ToR είναι ότι με χαμηλότερη κατανάλωση ενέργειας μπορεί να διαχειριστεί εύκολα 10 Tbps κίνησης (400 Gbps με 25 κανάλια flex-grid). Κάθε rack μπορεί να συνδέσει διακομιστές με πολλές χιλιάδες πυρήνες μέσω καρτών NIC υψηλού ρυθμού δεδομένων. Ο μεταγωγέας πυρήνας της αρχιτεκτονικής χρησιμοποιεί πολλές 1000×1000 MEMS και ένα flexi-grid WSS. Δημιουργείται ένα μη μπλοκαρισμένο δίκτυο Clos με τη χρήση πολλαπλών σταδίων MEMS διακοπών. Αυτός ο κεντρικός switch διευκολύνει την επικοινωνία μεταξύ των ραφιών. Ο χρόνος εγκατάστασης των MEMS διαρκεί μερικά χιλιοστά του δευτερολέπτου. Ωστόσο, καθώς η επικοινωνία μεταξύ των ραφιών βασίζεται κυρίως στους MEMS, το δίκτυο μπορεί να είναι αργό όταν δεν υπάρχει ήδη κατάλληλο κύκλωμα.

Το μέγεθος των διακοπών MEMS εξαρτάται από τον αριθμό των χρησιμοποιούμενων καναλιών flex-grid. Συγκεκριμένα, ο switch πυρήνα απαιτεί f θύρες εισόδου, όπου f είναι ο αριθμός των ραφιών και f ο αριθμός των καναλιών flex-grid.

Για παράδειγμα, για την υποστήριξη 100 ραφιών και 100 κανάλια flex-grid, απαιτούνται 10.000 θύρες εισόδου. Αυτή η χωρητικότητα θυρών καλύπτεται με τον συνδυασμό πολλών μεγάλων διακοπών MEMS. Παρόμοια με τον switch Petabit, ο αριθμός των σταδίων m που απαιτούνται για το δίκτυο Clos εξαρτάται από τη σχέση $m = \log_{\gamma} f$, όπου γ είναι ο αριθμός των θυρών που υποστηρίζονται από έναν μόνο switch MEMS.

Το EODCN προτείνει τη χρήση χρονοπρογραμματιστών εντός και μεταξύ των σχαρών για τον υπολογισμό του χρόνου και της διάρκειας μετάδοσης των πακέτων, όπως φαίνεται στο Σχήμα 2.6β. Μέχρι να ολοκληρωθεί ο υπολογισμός της μετάδοσης, τα πακέτα αποθηκεύονται στον κόμβο υπολογισμού/αποθήκευσης (CSN). Ο CSN χρησιμοποιεί έναν πομποδέκτη μεταβλητού εύρους ζώνης (BVT) για να συντονίζεται στο κατάλληλο μήκος κύματος και να υποστηρίζει πολλαπλούς ρυθμούς δεδομένων. Όταν ένα πακέτο μεταδίδεται από τον CSN μέσω του πομποδέκτη, το υπόλοιπο κύκλωμα είναι προκαθορισμένο. Έτσι, μετά την καθυστέρηση διάδοσης, ο αντίστοιχος δέκτης μπορεί να λάβει το πακέτο. Δεν υπάρχει προσωρινή αποθήκευση στο δίκτυο. Όταν απαιτείται η ρύθμιση κυκλωμάτων μεταξύ των πλαισίων, οι απομονωτές στον CSN πρέπει να είναι αρκετά μεγάλοι ώστε να κρατούν τα πακέτα μέχρι να είναι έτοιμο το αντίστοιχο μονοπάτι του δικτύου.

Ο πομποδέκτης (transceiver) στο CSN αποδεικνύεται ότι καταναλώνει 11 W ισχύος, γεγονός που συνιστά έναν σημαντικό περιοριστικό παράγοντα. Ο κεντρικός έλεγχος είναι παρόμοιος με αυτόν του OpenFlow, αλλά υπάρχουν δύο κύριες διαφορές:

- 1) Οι χρονοπρογραμματιστές υπολογίζουν τον χρόνο και τη διάρκεια μετάδοσης με συγκεντρωτικό τρόπο
- 2) Οι χρονοπρογραμματιστές υπολογίζουν το χρονοδιάγραμμα για πακέτα ή ριπές πακέτων και όχι για ροές

Επομένως, η κυκλοφορία στο επίπεδο ελέγχου είναι πιο αυξημένη σε σχέση με τους χρονοπρογραμματιστές του OpenFlow. Για το λόγο αυτό, αυτοί οι χρονοπρογραμματιστές έχουν προταθεί για χρήση σε προγραμματιζόμενες συστοιχίες πυλών πεδίου, προκειμένου να μειωθεί ο χρόνος χρονοπρογραμματισμού.

Συνολικά, η αρχιτεκτονική είναι απλή, μειώνει την κατανάλωση ενέργειας σε όλο το δίκτυο και προσφέρει επεκτασιμότητα. Ωστόσο, τα κύρια μειονεκτήματα περιλαμβάνουν την κατανάλωση ισχύος του πομποδέκτη CSN και την επίδραση στην απόδοση των επικοινωνιών μεταξύ των ραφιών λόγω του χρόνου εγκατάστασης του κυκλώματος. Οι οπτικές απώλειες περιορίζουν τον αριθμό των CSN που μπορούν να υποστηριχτούν ανά ράφι από τον οπτικό μεταγωγέα ToR. Με περιθώριο ισχύος 30 dB, το δίκτυο μπορεί να υποστηρίξει λιγότερα από 20 CSN ανά rack.

2.3.6 Αρχιτεκτονική OGDCN

Το δίκτυο κέντρων δεδομένων με οπτική υποδομή (OGDCN), παρουσιάζει μια οπτική αρχιτεκτονική DCN που είναι συγκρίσιμη με το EODCN. Περιλαμβάνει οπτικούς διαχωριστές-συνδυαστές, αναγεννητές σήματος και WSS σε όλους τους μεταγωγείς του. Χάρη στους αναγεννητές σήματος, ένας switch ToR στο OGDCN μπορεί να υποστηρίξει μεγαλύτερο λόγο fan-out σε σύγκριση με το αντίστοιχο EODCN. Η βασική διαφορά είναι η συνεχής διαθεσιμότητα της συνδεσιμότητας any-to-any CSN στο OGDCN. Δεν απαιτούνται επαναδιαμορφώσεις του μεταγωγέα για τη δημιουργία μιας διαδρομής δικτύου από άκρο σε άκρο. Επομένως, το OGDCN είναι ένα από τα πιο σημαντικά συστήματα που μπορούν να χρησιμοποιηθούν, μειώνοντας σημαντικά τις απαιτήσεις απομονωτικού χώρου στο CSN. Το CSN σε μια αρχιτεκτονική OGDCN δεν χρειάζεται μεγάλα πακέτα απομονωτών πακέτων.

Όσον αφορά τη λειτουργία του δικτύου, το OGDCN χρησιμοποιεί ένα σύνολο κεντρικών ελεγκτών για τη διαχείριση της μετάδοσης. Οι υπόλοιπες λειτουργίες ελέγχου της διαδρομής είναι παρόμοιες με αυτές του EODCN (βλέπε Σχήμα 12).

2.3.7 Αρχιτεκτονική OSA

Η αρχιτεκτονική OSA, ενσωματώνει υβριδικούς switch. Αυτή η αρχιτεκτονική χρησιμοποιεί ένα ηλεκτρονικό SwF και μια σειρά πομποδεκτών, τοποθετώντας WSS μπροστά από τους πομποδέκτες διασύνδεσης. Βασίζεται σε μια τοπολογία δικτύου Torus με δύο ή περισσότερες διαστάσεις. Σε κάθε διάσταση, υπάρχουν δύο γείτονες, με αποτέλεσμα 2N γείτονες σε μια τοπολογία Torus N διαστάσεων. Το WSS αντιστοιχίζει ένα εξερχόμενο μήκος κύματος σε έναν από τους γείτονές του.

Η λειτουργία του δικτύου μοιάζει με αυτή ενός μεταγωγέα πακέτων, καθώς εκτός από τον υπολογισμό μιας θύρας εξόδου για κάθε πακέτο, υπολογίζεται επίσης και ένα μήκος κύματος εξόδου, όπως αναφέρεται παρακάτω βήμα προς βήμα:

1) Ο πυρήνας διαθέτει συνήθως πολλές διαδρομές προς κάθε προορισμό. Μία από αυτές τις διαδρομές επιλέγεται για την τρέχουσα ροή και η αντίστοιχη εξερχόμενη διεπαφή επιλέγεται για τη μετάδοση των δεδομένων.

2) Σε αυτό το σημείο, εντοπίζεται ένα διαθέσιμο μήκος κύματος που φτάνει σε αυτόν τον γείτονα.

3) Το πακέτο μεταδίδεται μέσω αυτού του μήκους κύματος στον γείτόνα του.

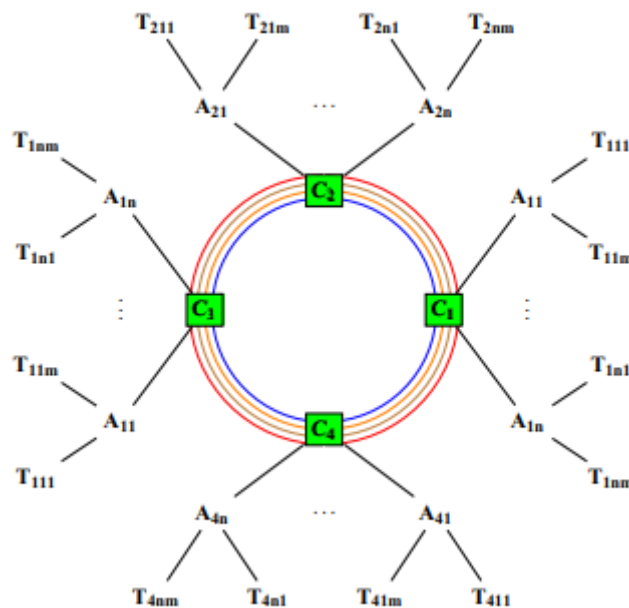
4) Στον γείτονα, το μήκος κύματος αυτό μετατρέπεται από ένα WSS σε έναν δέκτη που αποτελεί μέρος της συστοιχίας δεκτών. Ο δέκτης αυτός αποκωδικοποιεί το πακέτο και υπολογίζει ξανά το επόμενο βήμα, ακολουθώντας τα προηγούμενα βήματα, μέχρι το πακέτο να φτάσει στον τελικό προορισμό.

Σε αυτή την αρχιτεκτονική, όλα τα πακέτα επεξεργάζονται σε κάθε άλμα κατά τη διάρκεια της διαδρομής τους μέσα στο δίκτυο. Αφού καθοριστεί το επόμενο βήμα και το μήκος κύματος για κάθε πακέτο, ο μεταγωγέας χρησιμοποιεί αυτές τις πληροφορίες για να προχωρήσει την αποστολή του πακέτου στο επόμενο βήμα. Για τη δρομολόγηση των πακέτων χρησιμοποιείται ένας απλός μηχανισμός αναζήτησης, παρόμοιος με τη δρομολόγηση XY.

Σε κάθε άλμα, όλα τα εισερχόμενα μήκη κύματος μετατρέπονται ξανά σε πακέτα. Έτσι, τα πακέτα που προορίζονται για έναν τοπικό υπολογιστικό διακομιστή αντιγράφονται τοπικά, ενώ τα υπόλοιπα πακέτα προωθούνται. Αυτή η αρχιτεκτονική επιτρέπει την αποδοτική αξιοποίηση των διαθέσιμων θέσεων σε μήκη κύματος, γεμίζοντας τυχόν κενές θέσεις.

2.3.8 FISSION

Η αρχιτεκτονική FISSION (Flexible Interconnect for Scalable Integrated Systems using Optical Networks), αποσκοπεί κυρίως στην ενίσχυση της απόδοσης του δικτύου στο επίπεδο πυρήνα μιας ιεραρχικής τοπολογίας, όπως απεικονίζεται στο Σχήμα 2.7. Αυτή η αρχιτεκτονική χρησιμοποιεί έναν υβριδικό μεταγωγέα για τη βελτιστοποίηση της απόδοσης του δικτύου.



Σχήμα 2.7 Τοπολογία μιας αρχιτεκτονικής FISSION[2]

Για την επίτευξη αυτού του στόχου, η αρχιτεκτονική εισάγει έναν ηλεκτρο-οπτικό switch που μπορεί να συνδεθεί τόσο με ηλεκτρονικά όσο και με οπτικά δίκτυα. Στον οπτικό τομέα,

χρησιμοποιείται ένας αναδιαμορφώσιμος πολυπλέκτης προσθήκης/πτώσης (ROADM). Κάθε κεντρικός switch συμμετέχει σε ένα σύνολο δακτυλίων οπτικών ινών που συνδέουν άλλους μεταγωγείς πυρήνα. Αυτοί οι δακτύλιοι παρέχουν αμφίδρομη επικοινωνία μεταξύ των μεταγωγέων πυρήνα μέσω του οπτικού τομέα. Με αυτή τη διάταξη, κάθε κύκλωμα μπορεί να χρησιμοποιηθεί για την επικοινωνία από μια πηγή σε πολλούς προορισμούς, επιτρέποντας έτσι την αποτελεσματική μετάδοση δεδομένων μεταξύ των κεντρικών μεταγωγέων.

Το δίκτυο λειτουργεί δημιουργώντας ένα κύκλωμα end-to-end μεταξύ ενός συνόλου κεντρικών μεταγωγέων, το οποίο χρησιμοποιείται για τη μεταφορά δεδομένων. Το μήκος κύματος που αντιστοιχεί σε αυτό το κύκλωμα προστίθεται από τον κεντρικό μεταγωγέα προέλευσης και χρησιμοποιεί ένα συγκεκριμένο δακτύλιο οπτικών ινών για τη σύνδεση. Αυτό το μήκος κύματος μπορεί να απορριφθεί από έναν ή περισσότερους μεταγωγείς πυρήνα.

Η προώθηση των πακέτων βασίζεται σε μια στρατηγική εκπομπής και επιλογής, όπου οι ενδιαμέσοι μεταγωγείς πυρήνα παρακάμπτουν τα πακέτα αν δεν έχουν ενδιαφέρον γι' αυτά. Σε ό,τι αφορά τη λειτουργία του δικτύου, η αρχιτεκτονική είναι παρόμοια με το OSA, με τη διαφορά ότι στο FISSION υπολογίζεται το εξερχόμενο κύκλωμα αντί για το εξερχόμενο μήκος κύματος.

Η τοπολογία του οπτικού πεδίου είναι πολυδιάστατη Torus, με τον αριθμό των δακτυλίων να ισούται με τις διαστάσεις του Torus και κάθε δακτύλιος να συνδέει έναν σταθερό αριθμό κόμβων.

Η αρχιτεκτονική FISSION διαφέρει από το OSA κυρίως σε δύο σημεία:

- 1) Εφαρμόζει οπτική παράκαμψη, μειώνοντας τις μετατροπές από οπτικό σε ηλεκτρονικό και κατά συνέπεια την κατανάλωση ενέργειας.
- 2) Χρησιμοποιεί οπτικούς δακτύλιους για τη σύνδεση μεταγωγέων πυρήνα και ToR, σε αντίθεση με το OSA.

2.4 Σύγκριση αρχιτεκτονικών

Στον Πίνακα 2.3 συνοψίζονται τα χαρακτηριστικά του υβριδικού switch σε διάφορες αρχιτεκτονικές, αναφέροντας την τοπολογία, τα στοιχεία διασύνδεσης, και τα συστατικά όπως SwF, ρυθμιστές θύρας και μηχανισμούς αναζήτησης. Για παράδειγμα, η αρχιτεκτονική Helios χρησιμοποιεί ιεραρχική τοπολογία, συνδετήρες για διασύνδεση και MEMS για μεταγωγή, χωρίς θύρες απομονωτές ή μηχανισμό αναζήτησης. Από την άλλη, η αρχιτεκτονική FISSION βασίζεται σε δακτύλιο, χρησιμοποιεί ROADMs για διασύνδεση και οπτικούς δακτυλίους για μεταγωγή.

Architecture ↓	Topology	Interface component	Switch fabric	Port buffers	Lookup
Fat-tree	Hierarchical	Receiver and transmitter	Crossbar	Ingress and egress buffers	IP or label lookup
Fastpass	Hierarchical	Receiver and transmitter	Crossbar	None	Centralized scheduler
LION	Hierarchical	TWC and Channel adapter	AWGR	Delay line	Label lookup
TONAK-LION	Hierarchical	TWC and channel adapter	2 AWGR	Delay line	Label lookup
Petabit	Hierarchical	TWC, TLs and channel adapter	m-stage AWGR	Ingress buffers	Port matching
Helios	Hierarchical	Connectors	MEMS	None	None
EODCN ToR	Hierarchical	Connectors	WSS and splitter	None	Intra-rack scheduler
Core		Connectors	WSS & m-stage MEMS	None	Inter-rack scheduler
OGDCN	Hierarchical	Connectors	WSS, regenerators and splitter	None	Centralized scheduler
OSA	Torus	WSS and transceiver array	Crossbar	ingress and egress buffers	Simplified lookup
OSA-ND	>2 D-Torus	WSS and transceiver array	Crossbar	ingress and egress buffers	Simplified lookup
FISSION	Ring	ROADM	Fiber rings (no switching)	Ingress and egress buffers	Broadcast and select

Πίνακας 2.3 Σύγκριση αρχιτεκτονικών βάσει στοιχείων[2]

Η αρχιτεκτονική Helios μειώνει τον αριθμό των πομποδεκτών μεταξύ των ραφιών, αλλά η αργή προσαρμογή στις αλλαγές κίνησης λόγω του χρόνου εγκατάστασης των MEMS (χρόνου χιλιοστών του δευτερολέπτου) είναι ένα μειονέκτημα. Η αρχιτεκτονική LION, αν και προσφέρει γρήγορη διαμόρφωση με TWC σε νανοδευτερόλεπτα, έχει περιορισμούς επεκτασιμότητας. Η αρχιτεκτονική Petabit ξεπερνά αυτούς τους περιορισμούς, αλλά απαιτεί μεγάλες ηλεκτρονικές απομονώσεις για την αποθήκευση πακέτων μέχρι την επίλυση συγκρούσεων. Η αρχιτεκτονική OGDCN προσφέρει κλιμακούμενη λύση αλλά λείπουν χαρακτηριστικά αξιοπιστίας. Τέλος, οι αρχιτεκτονικές FISSION και OSA χρησιμοποιούν τοπολογία Torus (δακτύλιο) και εκμεταλλεύονται τις πολλαπλές διαδρομές για καλύτερη απόδοση.

Οι περισσότερες αρχιτεκτονικές χρησιμοποιούν μεταγωγείς πακέτων στο στρώμα ToR, και μερικές έχουν επιπλέον μεταγωγείς στο δίκτυό τους, κάτι που αυξάνει την κατανάλωση ενέργειας. Ωστόσο, τα EODCN και OGDCN προσφέρουν μεταγωγείς ToR με υψηλή αποδοτικότητα, καταναλώνοντας σημαντικά λιγότερη ενέργεια σε σύγκριση με τους μεταγωγείς πακέτων.

Σχετικά με την επεκτασιμότητα και την απόδοση, οι αρχιτεκτονικές FISSION, OSA, Petabit και OGDCN δείχνουν σημαντική προοπτική. Από ενεργειακή άποψη, οι Helios, EODCN και OGDCN είναι πιο αποδοτικές λόγω της μειωμένης χρήσης ηλεκτρονικών εξαρτημάτων. Αξιοπιστίας, οι OSA και FISSION επωφελούνται από την τοπολογία Torus με πλεονάζουσες διαδρομές δικτύου.

Δεν υπάρχει μία αρχιτεκτονική που να ξεχωρίζει σε όλους τους τομείς, αλλά με βάση τις ανάγκες, μπορεί να επιλεγεί η πιο κατάλληλη. Επιπλέον, η συνδυασμένη χρήση των καλύτερων εξαρτημάτων (components) από διάφορες αρχιτεκτονικές μπορεί να οδηγήσει στη δημιουργία νέων, πιο αποδοτικών και αξιόπιστων αρχιτεκτονικών DCN. Οι επιδόσεις των συστατικών σε όρους κατανάλωσης ενέργειας και καθυστέρησης είναι ήδη γνωστές, επιτρέποντας τη δημιουργία βελτιωμένων αρχιτεκτονικών που συνδυάζουν τις επιθυμητές χαρακτηριστικά.

Κεφάλαιο 3 : Βασικές αρχές τεχνητής νοημοσύνης και οι απαιτήσεις της

3.1 Εισαγωγή στην τεχνητή νοημοσύνη

Η τεχνητή νοημοσύνη (TN) αποτελεί έναν από τους πιο συναρπαστικούς και ραγδαία αναπτυσσόμενους τομείς της σύγχρονης τεχνολογίας, με την ικανότητα να αλλάζει ριζικά τον τρόπο που ζούμε και εργαζόμαστε. Η TN βρίσκει πεδίο εφαρμογής σε τομείς όπως η ιατρική, η εκπαίδευση, και η βιομηχανία. Μέσα από την αξιοποίηση πολύπλοκων αλγορίθμων και την επεξεργασία τεράστιων όγκων δεδομένων (big data), η τεχνητή νοημοσύνη μπορεί να επιτύχει αποτελέσματα που κάποτε θεωρούνταν αδύνατα για τις μηχανές, όπως η αναγνώριση εικόνων, η ανάλυση φυσικής γλώσσας και η λήψη αποφάσεων. Στον τομέα των δικτύων η TN δύναται να κατέχει καθοριστικό ρόλο στο κομμάτι της λήψης αποφάσεων σε διάφορες καταστάσεις όπως για παράδειγμα η επιλογή του καταλληλότερου μονοπατιού (path) για την μεταφορά δεδομένων από έναν κόμβο σε έναν άλλο. Στις παρακάτω υποενότητες θα αναλυθούν κάποιες από τις βασικές αρχές της τεχνητής νοημοσύνης αλλά και οι απαιτήσεις τις ώστε να εξασφαλίζεται η ομαλή λειτουργικότητά της.

3.2 Βασικές αρχές τεχνητής νοημοσύνης

Παρακάτω αναλύονται κάποιες από τις βασικές αρχές της TN, οι οποίες αποτελούν και τις θεμελιώδεις βάσεις της.

Μηχανική Μάθηση

Η Μηχανική Μάθηση (Machine Learning - ML) είναι ένας τομέας της τεχνητής νοημοσύνης (TN) που εστιάζει στην ανάπτυξη αλγορίθμων και τεχνικών που επιτρέπουν στα υπολογιστικά συστήματα να μαθαίνουν και να βελτιώνονται από την εμπειρία, χωρίς να προγραμματίζονται ρητά για κάθε ενέργεια. Η μηχανική μάθηση δίνει τη δυνατότητα στα συστήματα να προσαρμόζονται και να αναπτύσσουν τις ικανότητές τους μέσω ανάλυσης δεδομένων.

Αλγόριθμοι Μηχανικής Μάθησης

Επιβλεπόμενη Μάθηση (Supervised Learning): Το μοντέλο εκπαιδεύεται με δεδομένα που περιλαμβάνουν εισόδους και τις αντίστοιχες εξόδους (ετικέτες). Παραδείγματα περιλαμβάνουν τη γραμμική παλινδρόμηση και τις αλγορίθμους ταξινόμησης, όπως οι δέντρα αποφάσεων και οι υποστηρικτικοί διανυσματικοί μηχανισμοί (SVM). Στην επιβλεπόμενη μάθηση (SL), χρησιμοποιούνται μετρικές όπως η ακρίβεια, η ανάκληση, η F1-score και η μέση τετραγωνική απόκλιση (MSE) για την αξιολόγηση της απόδοσης του μοντέλου.

Μη επιβλεπόμενη Μάθηση (Unsupervised Learning): Το μοντέλο προσπαθεί να αναγνωρίσει μοτίβα ή δομές στα δεδομένα χωρίς προηγούμενες ετικέτες. Παραδείγματα περιλαμβάνουν την ανάλυση συστατικών (PCA) και τη συμπλέκτηση (clustering) μέσω αλγορίθμων όπως το K-means.

Μάθηση Ενίσχυσης (Reinforcement Learning): Το μοντέλο μαθαίνει μέσω αλληλεπίδρασης με το περιβάλλον και λαμβάνει ανατροφοδότηση σε μορφή επιβράβευσης ή ποινής. Το συγκεκριμένο μοντέλο χρησιμοποιείται ευρέως στα νευρωνικά δίκτυα (NN).

Η μηχανική μάθηση έχει ευρύ φάσμα εφαρμογών σε διάφορους τομείς:

Υγειονομική Περίθαλψη: Ανάλυση ιατρικών εικόνων, διάγνωση ασθενειών και προσωποποιημένη ιατρική.

Εμπόριο και Λιανική: Σύσταση προϊόντων, ανάλυση αγοραστικών προτιμήσεων και βελτιστοποίηση αποθεμάτων.

Μεταφορές: Αυτόνομα οχήματα, βελτιστοποίηση διαδρομών και ανάλυση της κίνησης στους δρόμους.

Η μηχανική μάθηση συνεχώς εξελίσσεται με την ανάπτυξη νέων αλγορίθμων και τεχνικών, ανοίγοντας νέους δρόμους για καινοτόμες εφαρμογές και την επίλυση σύνθετων προβλημάτων σε πολλούς τομείς.

Νευρωνικά δίκτυα (Deep Neural Networks) και βαθιά μάθηση(Deep Learning)

Η βαθιά μάθηση(DL) αποτελεί μία τεχνική μηχανικής μάθησης (ML) που χρησιμοποιεί νευρωνικά δίκτυα (NN) για να προσεγγίσει λειτουργίες που σχετίζονται με την ταξινόμηση, την παλινδρόμηση ή τη λήψη αποφάσεων. Το βασικό δομικό στοιχείο κάθε συστήματος DL είναι τα NN, τα οποία αποτελούν μαθηματικές δομές και χρησιμοποιούν γραμμικά σταθμισμένες προσθέσεις μαζί με μη γραμμικά φίλτρα για την επεξεργασία πολυδιάστατων εισόδων. Αυτά τα δίκτυα μπορούν να προσεγγίσουν αυθαίρετες λειτουργίες καθιστώντας τα εξαιρετικά χρήσιμα σε μια τεράστια ποικιλία λειτουργιών, όπως η ταξινόμηση εικόνας, η σύσταση περιεχομένου, η διαχείριση πόρων υπολογιστικών συστημάτων και πολλά άλλα.

Ένα νευρωνικό δίκτυο (NN) μπορεί να θεωρηθεί ως μια παραμετροποιημένη συνάρτηση

$$f(x, \theta) \rightarrow y,$$

όπου η συνάρτηση f λαμβάνει ως είσοδο έναν ταυυστή(tensor) x , παραμετροποιείται μέσω των παραμέτρων θ και παράγει ως έξοδο τον y . Τα νευρωνικά δίκτυα εκπαιδεύονται μέσω μιας αντικειμενικής συνάρτησης, η οποία προσαρμόζει τις παραμέτρους της συνάρτησης με βάση τη σύγκριση μεταξύ της πραγματικής και της επιθυμητής εξόδου.

Τα νευρωνικά δίκτυα (NN) είναι οργανωμένα σε στρώματα, όπου κάθε στρώμα αποτελείται από έναν αριθμό μονάδων. Κάθε μονάδα λαμβάνει ως είσοδο το σταθμισμένο άθροισμα των εξόδων των μονάδων του προηγούμενου στρώματος και παράγει μια έξοδο αφού περάσει αυτό το άθροισμα μέσα από μια μη γραμμική λειτουργία. Τα βαθιά νευρωνικά δίκτυα αναφέρονται απλώς σε δίκτυα που διαθέτουν πολλαπλά στρώματα, σε αντίθεση με αυτά που έχουν μόνο ένα στρώμα.

Πιο συγκεκριμένα, κάθε στρώμα, l , μπορεί να θεωρηθεί ως ένας ταυυστής με παραμέτρους :

$$W^l = \sum w_{ij}^l$$

Δεδομένου ενός n -διάστατου διανύσματος εισόδου x και ενός στρώματος εισόδου W^o με διαστάσεις $(M \times N)$ σε ένα νευρωνικό δίκτυο (NN), η έξοδος της k -οστής μονάδας του πρώτου στρώματος, h^o , υπολογίζεται μέσω του πολλαπλασιασμού της $W^o x$, και στη συνέχεια περνάει μέσα από μια μη γραμμική συνάρτηση ενεργοποίησης, $\sigma(x)$. Επιπλέον, ένας όρος πόλωσης b προστίθεται στο άθροισμα, παρόμοια με τον σταθερό όρο b στην τυπική γραμμική συνάρτηση $y = ax + b$:

$$h_k^l = \sigma(w_{k0}^l x_0 + \dots + w_{kN}^l x_N) = \sigma(b_k^l + \sum_{j=0}^N w_{kj}^l x_j)$$

και γενικότερα, για ένα NN με $n \geq 1$, η έξοδος του l στρώματος είναι :

$$h^l = \sigma(h^{l-1T} W^l + b^l)$$

Αξίζει να τονιστεί ότι τα βαθιά νευρωνικά δίκτυα έχουν κύριο ρόλο σε αρκετές εφαρμογές TN κατά τις οποίες γίνεται χρήση της ενισχυμένης μάθησης (Reinforcement Learning), όπως αναφέρθηκε και προηγουμένως για την πραγματοποίηση επιθυμητών στόχων.

Επεξεργασία Φυσικής Γλώσσας(Natural Language Processing -NPL)

Η επεξεργασία φυσικής γλώσσας, ή Natural Language Processing (NLP), είναι μια δημοφιλής εφαρμογή της τεχνητής νοημοσύνης (TN) που επικεντρώνεται στην αλληλεπίδραση ανθρώπου-υπολογιστή μέσω φυσικής γλώσσας. Παρά τα οφέλη της, όπως η δυνατότητα επικοινωνίας με υπολογιστές σε φυσικές γλώσσες και η ανάκτηση πληροφοριών με πιο φυσικό τρόπο, η NPL αντιμετωπίζει προκλήσεις λόγω της ασάφειας στη σύνταξη, τη σημασιολογία και το λεξιλόγιο.

Η διαδικασία της NPL περιλαμβάνει τρία βασικά στάδια:

- 1) **Συντακτική ανάλυση:** Επικεντρώνεται στη γραμματική δομή των προτάσεων.
- 2) **Σημασιολογική ανάλυση:** Μετατρέπει τις προτάσεις σε εσωτερικές αναπαραστάσεις βασισμένες στη σημασία τους.
- 3) **Πραγματολογική ανάλυση:** Εντάσσει τις προτάσεις σε ένα γενικό νοηματικό σύνολο, λαμβάνοντας υπόψη τις συνθήκες υπό τις οποίες εκφράστηκαν.

Η NPL προσφέρει ποικίλα οφέλη, όπως η φωνητική αναζήτηση, η αυτόματη μετάφραση κειμένων και η εξόρυξη δεδομένων, διευκολύνοντας την αλληλεπίδραση με υπολογιστές και την επεξεργασία μεγάλων κειμένων.

Μηχανική Όραση (Machine Vision)

Η Μηχανική Όραση, γνωστή και ως machine vision, αποτελεί ένα πεδίο της τεχνητής νοημοσύνης (TN) που ασχολείται με την ικανότητα των υπολογιστών να αντιλαμβάνονται και να ερμηνεύουν οπτικά στοιχεία από το περιβάλλον, όπως εικόνες και βίντεο. Μέσω εξελιγμένων αλγορίθμων και καινοτόμων μεθόδων μηχανικής μάθησης, η τεχνολογία αυτή δίνει τη δυνατότητα στους υπολογιστές να αναγνωρίζουν αντικείμενα, να εντοπίζουν πρόσωπα, να αναλύουν σκηνές και να κατανοούν πολύπλοκες οπτικές πληροφορίες για να εκτελούν καθήκοντα που απαιτούν οπτική αντίληψη. Οι εφαρμογές της περιλαμβάνουν την αυτόματη αναγνώριση προσώπων, την ανάλυση βίντεο ασφαλείας, την αυτόνομη οδήγηση οχημάτων, καθώς και την ιατρική διάγνωση μέσω εικόνων. Με την συνεχώς αυξανόμενη ακρίβεια και αποτελεσματικότητα των αλγορίθμων οπτικής αντίληψης, ο τομέας αυτός αναμένεται να έχει σημαντικό αντίκτυπο σε πολλούς τομείς, διευκολύνοντας τις αλληλεπιδράσεις με την τεχνολογία και ενισχύοντας την ποιότητα ζωής.

Αλγόριθμοι και Μαθηματικά Μοντέλα

Οι αλγόριθμοι και τα μαθηματικά μοντέλα είναι θεμελιώδη για την αποδοτικότερη λειτουργία των συστημάτων της ΤΝ. Κάποιοι από αυτούς του αλγόριθμους είναι οι:

Αλγόριθμοι Ενισχυτικής Μάθησης(RL)

Q-Learning

Η μέθοδος Q-Learning ανήκει σε μία κατηγορία αλγορίθμων στην ενισχυμένη μάθηση(RL), οι οποίοι επιτυγχάνουν την εκμάθηση της βέλτιστης πολιτικής έμμεσα, προσεγγίζοντας τη συνάρτηση Q (η οποία αντιπροσωπεύει την αξία κατάστασης-δράσης) για αυτή την πολιτική. Στη συνέχεια, η πολιτική μπορεί να εφαρμοστεί μέσω άπληστης επιλογής δράσεων, βασιζόμενη σε αυτή τη συνάρτηση Q. (π.χ $a = \max_{a^* \in A} Q(s, a^*)$))

Γενικά, αν η J είναι μια διαφορίσιμη συνάρτηση J(w), παραμετροποιημένη από τις παραμέτρους w, τότε η κλίση της έχει τη μορφή:

$$\nabla_{\mathbf{w}} J(\mathbf{w}) = \left[\frac{\partial J(\mathbf{w})}{\partial w_1}, \dots, \frac{\partial J(\mathbf{w})}{\partial w_n} \right]$$

Εάν η J(w) αντιπροσωπεύει μια αντικειμενική συνάρτηση που σχετίζεται με μια διαφορίσιμη συνάρτηση Q(s, a), τότε η παράγωγός της μπορεί να αξιοποιηθεί για την επαναληπτική βελτίωση της Q(s, a) προς κάποια βέλτιστη λύση, χρησιμοποιώντας τη μέθοδο της βαθμωτής καθόδου.

Για παράδειγμα, για πραγματικές και προσεγγιστικές συναρτήσεις Qπ και Q', αντίστοιχα:

$$J(\mathbf{w}) = \mathbb{E}[(Q_{\pi}(s, a) - Q'(s, a, \mathbf{w}))^2]$$

$$\Delta \mathbf{w} = \alpha (Q_{\pi}(s, a) - Q'(s, a, \mathbf{w})) \Delta_{\mathbf{w}} Q'(s, a, \mathbf{w})$$

Ένας περιορισμός της Q-Learning είναι ότι αξιολογεί τις ενέργειες χωρίς να τις καθορίζει άμεσα. Αυτό δημιουργεί δυσκολίες όταν ο χώρος των ενεργειών είναι πολύ μεγάλος ή συνεχής, καθώς η αξιολόγηση κάθε πιθανής ενέργειας μπορεί να γίνει δύσκολη. Σε τέτοιες περιπτώσεις, είναι απαραίτητο να χρησιμοποιηθούν πιο προσεγγιστικές μέθοδοι, όπως η προσέγγιση ενός συνεχούς χώρου αναζήτησης μέσω διακριτοποίησης.

Αλγόριθμοι Πολιτικής (Policy Gradient Methods)

Οι αλγόριθμοι πολιτικής μαθαίνουν απευθείας τις παραμέτρους για μια συνάρτηση πολιτικής. Παρόμοια με την εκμάθηση Q, μπορεί κανείς να ορίσει μια συνάρτηση στόχου $J(\theta)$, η οποία αξιολογεί την ποιότητα της τρέχουσας πολιτικής $\pi(\theta)$ με κάποιο τρόπο. Τόσο η συνάρτηση στόχου όσο και η πολιτική παραμετροποιούνται από τις παραμέτρους θ και είναι διαφορίσιμες.

Ομοίως με τα παραπάνω, ισχύουν :

$$\nabla_{\theta} J(\theta) = \left[\frac{\partial J(\theta)}{\partial \theta_1}, \dots, \frac{\partial J(\theta)}{\partial \theta_n} \right]$$

$$\Delta \theta = \alpha \nabla_{\theta} J(\theta)$$

Το $\Delta \theta$ αναφέρεται στο πόσο πρέπει να αλλάξει κάθε $(\theta_1, \dots, \theta_n)$ στο θ , και η κλίση $\nabla_{\theta} J(\theta)$ αναφέρεται σε αυτό το πλαίσιο ως η κλίση της πολιτικής. Μια γενική αναλυτική μορφή για την αντικειμενική συνάρτηση που χρησιμοποιείται για την αξιολόγηση μιας πολιτικής συνήθως σχετίζεται με την αναμενόμενη απόδοση που προκύπτει από την εφαρμογή αυτής της πολιτικής.

Η βαθμολογία συνάρτησης (score function), $\nabla_{\theta} \log(\pi_{\theta}(s, a))$, ορίζεται για λόγους ευκολίας από:

$$\nabla_{\theta} \pi_{\theta}(s, a) = \pi_{\theta}(s, a) \frac{\nabla_{\theta} \pi_{\theta}(s, a)}{\pi_{\theta}(s, a)} = \pi_{\theta}(s, a) \nabla_{\theta} \log(\pi_{\theta}(s, a))$$

Για ένα MDP ενός βήματος (one-step), εάν η αναμενόμενη απόδοση, σε σχέση με όλες τις πιθανές καταστάσεις, την πιθανότητα κάθε ενέργειας να πραγματοποιηθεί από την $\pi(s)$ και την σχετική ανταμοιβή, η αναμενόμενη ανταμοιβή είναι:

$$J(\theta) = \mathbb{E}_{\pi_{\theta}}[R] = \sum_{s \in \mathcal{S}} d(s) \sum_{a \in \mathcal{A}} \pi_{\theta}(s, a) R_{s,a}$$

επομένως ο αλγόριθμος πολιτικής είναι:

$$\nabla_{\theta} J(\theta) = \sum_{s \in \mathcal{S}} d(s) \sum_{a \in \mathcal{A}} \pi_{\theta}(s, a) \nabla_{\theta} \log(\pi_{\theta}(s, a)) \pi_{\theta}(s, a) R_{s,a} = \mathbb{E}_{\pi_{\theta}}[\nabla_{\theta} \log(\pi_{\theta}) R]$$

Όπου $d(s)$ αναφέρεται στην κατανομή των δειγμάτων που λαμβάνονται. Αυτή η συγκεκριμένη μορφή απαιτεί πλήρη επεισόδια για να μπορεί να υπολογίσει το συστατικό που σχετίζεται με την αναμενόμενη ανταμοιβή.

Δεδομένου ότι ο αλγόριθμος πολιτικής εξάγει ρητά τις δράσεις, οι μέθοδοί του μπορούν να παράγουν είτε πεπερασμένες (όπως κατηγορικές) είτε συνεχείς δράσεις. Επιπλέον, σε περιβάλλοντα όπου αυτό μπορεί να είναι χρήσιμο, όπως σε στοχαστικά περιβάλλοντα, οι αλγόριθμοι πολιτικής μπορούν να υλοποιήσουν στοχαστικές πολιτικές, σε αντίθεση με τις ντετερμινιστικές πολιτικές, όπως της εκμάθησης Q που περιγράφηκε παραπάνω.

Γενετικοί Αλγόριθμοι(Genetic Algorithms-GA)

Οι γενετικοί αλγόριθμοι (ΓΑ) ανήκουν στην κατηγορία των εξελικτικών αλγορίθμων. Λειτουργούν αναζητώντας λύσεις σε προβλήματα αρχικοποιώντας μια ομάδα υποψήφιων λύσεων και στη συνέχεια εξελίσσοντας επαναληπτικά αυτές τις λύσεις βάσει της ποιότητας του πληθυσμού, μέχρι να βρεθεί μια καλύτερη συγκλίνουσα λύση. Αυτές οι μέθοδοι είναι εμπνευσμένες από τη διαδικασία της εξέλιξης του DNA σε πληθυσμούς ζωντανών οργανισμών.

Οι γενετικοί αλγόριθμοι (ΓΑ) ξεκινούν με έναν αρχικό πληθυσμό ατόμων $g \in G$, όπου κάθε άτομο είναι μια υποψήφια λύση που συνήθως δημιουργείται τυχαία. Συνήθως, μια λύση αναπαρίσταται με μια συμβολοσειρά ή κάποιο ισοδύναμο σχήμα (όπως ένας πίνακας), που θυμίζει τις συμβολοσειρές χαρακτήρων που αποτελούν τα γονίδια στο DNA. Σε πολλά προβλήματα, αυτή η αναπαράσταση μπορεί να γίνει με τη μορφή πίνακα, είτε με διακριτές είτε με συνεχείς αριθμητικές τιμές. Τα άτομα αναφέρονται ως $g = [g_0, \dots, g_n]$ αν οι πληθυσμοί έχουν n χαρακτηριστικά ή διαστάσεις.

Πρέπει επίσης να οριστεί μια συνάρτηση καταλληλότητας $f : G \rightarrow \mathbb{R}$, η οποία αντιστοιχεί τα άτομα σε κλιμακωτές τιμές, με την προϋπόθεση ότι υψηλότερες τιμές καταλληλότητας αντιστοιχούν σε καλύτερες υποψήφιες λύσεις. Οι γενετικοί αλγόριθμοι εξελίσσουν τον πληθυσμό λύσεων επαναλαμβάνοντας τις διαδικασίες αξιολόγησης

καταλληλότητας \rightarrow επιλογής \rightarrow διασταύρωσης \rightarrow μετάλλαξης.

Η αξιολόγηση καταλληλότητας είναι η διαδικασία υπολογισμού της τιμής καταλληλότητας για κάθε μέλος του τρέχοντος πληθυσμού. Στη συνέχεια, η διαδικασία επιλογής δημιουργεί ένα νέο σύνολο από τον πληθυσμό με πιθανοτικό τρόπο. Αυτό συνήθως βασίζεται στην καταλληλότητα των ατόμων, όπου υψηλότερη καταλληλότητα σημαίνει μεγαλύτερη πιθανότητα να επιλεγούν για τη νέα ομάδα. Η δειγματοληψία γίνεται συχνά με αντικατάσταση από τον αρχικό πληθυσμό, έτσι ώστε τα άτομα με υψηλότερη καταλληλότητα να έχουν μεγαλύτερη πιθανότητα να εμφανιστούν πολλές φορές στο επιλεγμένο υποσύνολο.

Η διασταύρωση τυχαίων δειγμάτων περιλαμβάνει τη δημιουργία ζευγών ατόμων από το επιλεγμένο σύνολο και την εφαρμογή μερικής διασταύρωσης των γονιδίων τους. Συνήθως, αυτό γίνεται με τα εξής βήματα:

- 1) Επιλέγεται τυχαία ένα ζεύγος ατόμων από το επιλεγμένο σύνολο.
- 2) Επιλέγεται τυχαία ένα σημείο διασταύρωσης (δηλαδή ένα στοιχείο στο σύνολο $\{1, \dots, n\}$, όπου n είναι ο αριθμός των διαστάσεων ή γονιδίων).
- 3) Ανταλλάσσονται οι υποσειρές γονιδίων των δύο ατόμων που ακολουθούν το σημείο διασταύρωσης.

Αυτός ο τρόπος διασταύρωσης επιτρέπει τη δημιουργία νέων ατόμων που συνδυάζουν στοιχεία από τα γονίδια των δύο γονέων.

Ηθική και ασφάλεια

Η πρόοδος και η εφαρμογή της τεχνητής νοημοσύνης (TN) επιφέρει σημαντικά ηθικά και ασφαλιστικά ζητήματα που αγγίζουν σχεδόν όλους τους τομείς της ανθρώπινης δραστηριότητας. Αυτά τα ζητήματα περιλαμβάνουν την ευθύνη για τις αποφάσεις που λαμβάνονται από τα συστήματα TN, την προστασία της προσωπικής ζωής, την εξασφάλιση της διαφάνειας και την αποφυγή προκαταλήψεων.

Η προστασία της ιδιωτικότητας αποτελεί άλλη μια κρίσιμη διάσταση της ηθικής χρήσης της τεχνητής νοημοσύνης. Η συγκέντρωση και ανάλυση μεγάλων όγκων δεδομένων μπορεί να αποκαλύψει ευαίσθητες προσωπικές πληροφορίες. Η χρήση αυτών των δεδομένων χωρίς τη συγκατάθεση των χρηστών ή η αποθήκευσή τους χωρίς σαφείς πολιτικές προστασίας συνιστούν ανησυχητικές πρακτικές. Η εφαρμογή μέτρων όπως η μυστικότητα των δεδομένων και η ενσωμάτωση μηχανισμών για τη συγκατάθεση των χρηστών μπορεί να συμβάλει στην προστασία της ιδιωτικότητάς τους.

Η διαφάνεια των αλγορίθμων είναι επίσης θεμελιώδης. Οι χρήστες πρέπει να κατανοούν πώς λειτουργούν τα συστήματα TN και πώς λαμβάνονται οι αποφάσεις. Η έλλειψη διαφάνειας μπορεί να οδηγήσει σε απώλεια εμπιστοσύνης και σε αρνητικές συνέπειες. Τα συστήματα TN πρέπει να σχεδιάζονται με τρόπο που επιτρέπει την παρακολούθηση και ανασκόπηση των αποφάσεών τους.

Τέλος, τα συστημάτων TN πρέπει να προστατεύονται από κακόβουλες επιθέσεις και στην αποφυγή πιθανών κινδύνων που προκύπτουν από τις αδυναμίες τους. Για την ανάπτυξη ανθεκτικών και ασφαλών συστημάτων TN, είναι απαραίτητη η συνεχής αναβάθμιση της ασφάλειας και η εφαρμογή προληπτικών μέτρων για την αποτροπή επιθέσεων που μπορεί να εκμεταλλευτούν τις τυχόν αδυναμίες. Επίσης, η αποτελεσματική διαχείριση των κινδύνων απαιτεί την υιοθέτηση στρατηγικών ελέγχου και περιορισμού των πιθανών αρνητικών επιπτώσεων που μπορεί να προκύψουν από τη χρήση της TN.

3.3 Απαιτήσεις της Τεχνητής Νοημοσύνης

Παρακάτω θα αναλυθούν αρκετές από τις κύριες απαιτήσεις της TN, ώστε όλες οι διαδικασίες και οι μέθοδοι που λειτουργούν στα συστήματα TN, να κρίνονται ανταποκρίσιμοι και αποδοτικοί σε πραγματικές συνθήκες επίλυσης καθημερινών προβλημάτων που ενδέχεται να προκύπτουν κατά την εκτέλεσή τους. Συνεπώς οι συγκεκριμένες απαιτήσεις πρέπει να καλύπτονται ανά πάσα χρονική στιγμή.

Δεδομένα Μεγάλης Κλίμακας (Big Data/Big Flows)

Τα μεγάλα δεδομένα συχνά περιγράφονται από :

- 1) τον όγκο τους (δηλαδή το μέγεθός τους),
- 2) την ταχύτητα (δηλαδή την ταχύτητα απόκτησής τους)
- 3) την ποικιλία τους.

Αυτό σημαίνει συχνά ότι τα δεδομένα είναι πολυδιάστατα, περιλαμβάνοντας εικόνες, κείμενα και δομημένα πεδία. Η αυξανόμενη χρήση τους, έχει επιτρέψει τη συγκέντρωση δεδομένων από διάφορες πηγές, δημιουργώντας ένα πλούσιο περιβάλλον για μεγάλα δεδομένα και την τεχνητή νοημοσύνη.

Η διεξαγωγή έρευνας σε μεγάλα δεδομένα για την τεχνητή νοημοσύνη απαιτεί προσεκτική διαχείριση των ετικετών (labels) με τις οποίες εκπαιδεύονται τα μοντέλα, της δομής και των χαρακτηριστικών των δεδομένων, καθώς και των λεπτομερειών της αρχιτεκτονικής των μοντέλων τεχνητής νοημοσύνης. Είναι απαραίτητη η συμφωνία σχετικά με τους ορισμούς για την επισήμανση των δεδομένων, η υιοθέτηση προτύπων για την ανταλλαγή και επαναχρησιμοποίηση δεδομένων, η ανταλλαγή κώδικα που περιγράφει τα μοντέλα τεχνητής νοημοσύνης και η χρήση ανοιχτών API για τα μοντέλα αυτά.

Θεωρείται κρίσιμο να υπάρχουν κοινές ετικέτες δεδομένων, να υπάρχει η αποδοχή ως προς την ανταλλαγή δεδομένων και να είναι διαθέσιμα τα μοντέλα τεχνητής νοημοσύνης, ώστε να μεγιστοποιηθούν τα οφέλη από την έρευνα στον τομέα της τεχνητής νοημοσύνης.

Για να αξιοποιηθεί πλήρως το δυναμικό των μεγάλων δεδομένων και της τεχνητής νοημοσύνης, υπάρχουν αρκετοί κρίσιμοι τομείς που πρέπει να αντιμετωπιστούν στο μέλλον. Ένα βασικό κριτήριο είναι η δυνατότητα παραγωγής αναπαραγωγίμων αποτελεσμάτων. Όταν εκπαιδεύονται μοντέλα τεχνητής νοημοσύνης με μεγάλα δεδομένα, η αναπαραγωγιμότητα εξαρτάται σημαντικά από διάφορους παράγοντες, όπως ο τρόπος με τον οποίο χρησιμοποιούνται τα δεδομένα για την εκπαίδευση, η δομή και τα χαρακτηριστικά των δεδομένων, καθώς και οι λεπτομέρειες της αρχιτεκτονικής του μοντέλου. Αυτά τα στοιχεία είναι κρίσιμα και πρέπει να εξεταστούν προσεκτικά κατά την ανάπτυξη και την αξιολόγηση μοντέλων τεχνητής νοημοσύνης που βασίζονται σε μεγάλα δεδομένα.

Υπολογιστική Ισχύς

Η μηχανική μάθηση (ML) είναι μια υπολογιστικά απαιτητική διαδικασία, άρρηκτα συνδεδεμένη με την υπολογιστική ισχύ. Η υπολογιστική ισχύς και οι αρχιτεκτονικές των υπολογιστών επηρεάζουν την ταχύτητα εκπαίδευσης και εξαγωγής συμπερασμάτων στη μηχανική μάθηση, και κατ' επέκταση, τον ρυθμό προόδου της τεχνολογίας. Ωστόσο, οι σχέσεις αυτές είναι πιο περίπλοκες: το υλικό δεν επηρεάζει μόνο την ταχύτητα, αλλά και τις μεθόδους που χρησιμοποιούν οι ερευνητές και οι μηχανικοί για να σχεδιάσουν και να αναπτύξουν μοντέλα μηχανικής μάθησης.

Μια σημαντική παράμετρος είναι ότι το υλικό που χρησιμοποιείται για την εκπαίδευση ενός μοντέλου μηχανικής μάθησης μπορεί να διαφέρει σημαντικά από το υλικό που χρησιμοποιείται για την εξαγωγή συμπερασμάτων με ένα ήδη εκπαιδευμένο μοντέλο. Αυτό οφείλεται στις διαφορετικές απαιτήσεις που υπάρχουν σε κάθε στάδιο της ροής εργασίας της μηχανικής μάθησης. Για παράδειγμα, η κατανάλωση ενέργειας μπορεί να είναι κρίσιμη για ένα σύστημα υπολογιστικής όρασης που λειτουργεί σε μια κινητή συσκευή, ενώ μπορεί να μην είναι τόσο σημαντική όταν το σύστημα εκπαιδεύεται αρχικά σε ένα κέντρο δεδομένων.

Επιπλέον, η χρονική καθυστέρηση (Latency) μεταξύ της εισόδου και της εξόδου ενός συστήματος μπορεί να είναι καθοριστική σε ένα περιβάλλον υψηλής ταχύτητας, όπως η πλοήγηση, όπου η ταχύτητα εξαγωγής συμπερασμάτων επηρεάζει τον χρόνο αντίδρασης και διόρθωσης της πορείας. Όπως συμβαίνει με την κατανάλωση ενέργειας, η χρονική καθυστέρηση μπορεί να μην είναι τόσο σημαντική κατά την εκπαίδευση του συστήματος πλοήγησης. Αυτοί οι παράγοντες καθορίζουν ποιο είδος υλικού είναι κατάλληλο σε κάθε στάδιο του κύκλου ζωής ενός συστήματος μηχανικής μάθησης.

Στο πλαίσιο της υπολογιστικής ισχύς θα πρέπει εξίσου να αναφερθεί και η ανάλογη απαίτηση αποθηκευτικού χώρου με σκοπό την αποθήκευση των δεδομένων που συλλέγονται. Για παράδειγμα τα δεδομένα μεγάλης κλίμακας που αναφέρθηκαν στην προηγούμενη υποενότητα απαιτούν τον ανάλογο αποθηκευτικό χώρο από πλευράς υλικού αλλά ταυτόχρονα και την αντίστοιχη ικανότητα διαχείρισής τους.

Ποιότητα Υπηρεσίας(Quality Of Service – QoS)

Το Quality of Service (QoS) αναφέρεται σε ένα σύνολο τεχνολογιών και πρακτικών που αποσκοπούν στη διασφάλιση της βέλτιστης λειτουργίας και ποιότητας υπηρεσιών σε δίκτυα επικοινωνιών. Συγκεκριμένα, το QoS επικεντρώνεται στη διαχείριση της κίνησης του δικτύου ώστε να εξασφαλίζεται ότι συγκεκριμένες εφαρμογές ή υπηρεσίες, όπως η φωνή, το βίντεο και τα δεδομένα, λαμβάνουν την απαραίτητη προτεραιότητα για να διατηρούν υψηλή ποιότητα λειτουργίας, ακόμη και υπό συνθήκες αυξημένης ζήτησης στο δίκτυο.

Τα βασικά χαρακτηριστικά του QoS περιλαμβάνουν:

1) Καθυστέρηση (Delay): Ο χρόνος που απαιτείται για τη μετάδοση ενός πακέτου δεδομένων από το σημείο αποστολής στον προορισμό του. Το QoS μειώνει αυτήν την καθυστέρηση, κάτι που είναι κρίσιμο για εφαρμογές που απαιτούν άμεση απόκριση, όπως οι κλήσεις VoIP ή το streaming βίντεο.

2) Jitter: Η αστάθεια στην καθυστέρηση κατά την αποστολή πακέτων δεδομένων. Το QoS μπορεί να περιορίσει το jitter, βελτιώνοντας την ποιότητα υπηρεσιών όπως το βίντεο και η φωνή.

3) Απώλεια Πακέτων (Packet Loss): Η απώλεια δεδομένων κατά τη διάρκεια της μετάδοσής τους στο δίκτυο. Το QoS επιδιώκει να ελαχιστοποιήσει την απώλεια πακέτων, εξασφαλίζοντας έτσι πιο αξιόπιστη επικοινωνία.

Οι τεχνικές QoS εφαρμόζονται κυρίως σε δίκτυα IP, εταιρικά δίκτυα και δίκτυα κινητής τηλεφωνίας, εξασφαλίζοντας την απαραίτητη ποιότητα και αξιοπιστία σε υπηρεσίες που απαιτούν υψηλή απόδοση.

Παράμετροι QoS :

Όσον αφορά την καθυστέρηση,έχουμε:

$$\text{Delay} = R_i - S_i$$

όπου R_i και S_i είναι οι χρόνοι άφιξης και αποστολής του πακέτου i , αντίστοιχα.

Για το Jitter ισχύει ότι:

$$\text{Jitter} = |D_i - D_{i-1}|$$

όπου D_i και D_{i-1} είναι οι καθυστερήσεις για δύο διαδοχικά πακέτα.

Η απώλεια πακέτων(packet loss) είναι το ποσοστό(ratio) των μεταδιδόμενων πακέτων που δεν φθάνουν στον προορισμό τους, το οποίο υπολογίζεται ως εξής:

$$\text{Packet loss ratio} = \left(1 - \frac{\sum R_i(t)}{\sum S_i(t)}\right) \times 100$$

όπου $\sum R_i(t)$ και $\sum S_i(t)$ είναι ο συνολικός αριθμός των ληφθέντων και των μεταδοθέντων πακέτων κατά τη διάρκεια του i -οστού χρονικού διαστήματος, αντίστοιχα

Εκτίμηση και Αξιολόγηση Αλγορίθμων

Η εκτίμηση και η αξιολόγηση είναι κρίσιμα στάδια στη διαδικασία ανάπτυξης και διαχείρισης μοντέλων, είτε πρόκειται για αλγορίθμους μηχανικής μάθησης είτε για άλλου είδους μοντέλα. Αυτή η διαδικασία περιλαμβάνει δύο κύρια βήματα:

1)την αξιολόγηση της απόδοσης

και

2)την ανατροφοδότηση για συνεχή βελτίωση

Η αξιολόγηση της απόδοσης περιλαμβάνει την τακτική παρακολούθηση και αξιολόγηση της λειτουργικότητας των μοντέλων, με σκοπό την επιβεβαίωση ότι λειτουργούν όπως αναμένεται και την αναγνώριση αδυναμιών ή τομέων που χρειάζονται βελτίωση. Χρησιμοποιούνται διάφορες μετρήσεις και εργαλεία αξιολόγησης για να αναλυθεί η απόδοση των μοντέλων υπό πραγματικές συνθήκες.

Η ανατροφοδότηση που συλλέγεται επιτρέπει την πραγματοποίηση αλλαγών στο μοντέλο, όπως η προσθήκη νέων δεδομένων ή η τροποποίηση παραμέτρων. Για παράδειγμα, εάν το μοντέλο εντοπίζει λάθη ή έχει περιορισμένη αποτελεσματικότητα σε ορισμένα σενάρια, μπορεί να χρειαστεί να εκπαιδευτεί με περισσότερα ή καλύτερα δεδομένα, ή να γίνουν αλλαγές στις χαρακτηριστικές του παραμέτρους.

Η διαδικασία βελτίωσης είναι διαρκής και απαιτεί συστηματική παρακολούθηση και αναγνώριση των παραγόντων που επηρεάζουν την απόδοση του μοντέλου. Μπορεί να περιλαμβάνει πειραματισμούς με διάφορες τεχνικές ή αλγορίθμους, αναθεώρηση των δεδομένων εισόδου, ή ακόμη και αλλαγές στη δομή του μοντέλου.

Συνολικά, η εκτίμηση και αξιολόγηση μέσω της συνεχούς παρακολούθησης της απόδοσης και της ανατροφοδότησης είναι βασικά στοιχεία για την επιτυχία των μοντέλων. Εξασφαλίζουν ότι τα μοντέλα παραμένουν χρήσιμα, αξιόπιστα και ικανά να προσαρμόζονται σε νέες προκλήσεις και απαιτήσεις.

3.4.1 Συνελικτικά Νευρωνικά Δίκτυα(Convolutional Neural Networks – CNNs)

Ένα συνελικτικό νευρωνικό δίκτυο (CNN) είναι ένας τύπος μοντέλου μηχανικής μάθησης(ML), πιο συγκεκριμένα ένας αλγόριθμος βαθιάς μάθησης(DL) που είναι ιδανικός για την επεξεργασία οπτικών δεδομένων. Τα CNN, γνωστά και ως convnets, εφαρμόζουν αρχές από τη γραμμική άλγεβρα, όπως οι πράξεις συνέλιξης, για να εξάγουν χαρακτηριστικά και να αναγνωρίζουν μοτίβα σε εικόνες. Παρόλο που τα CNN χρησιμοποιούνται κυρίως για ανάλυση εικόνων, μπορούν επίσης να προσαρμοστούν για να επεξεργάζονται ηχητικά δεδομένα και άλλα είδη σημάτων.

Τα CNN αποτελούνται από μια αλληλουχία επιπέδων(layers), όπου το καθένα εντοπίζει διαφορετικά χαρακτηριστικά της εικόνας που εισάγεται. Ανάλογα με την πολυπλοκότητα του στόχου του, ένα CNN μπορεί να έχει δεκάδες, εκατοντάδες, ή ακόμα και χιλιάδες στρώματα, με κάθε επίπεδο να χρησιμοποιεί τα αποτελέσματα των προηγούμενων επιπέδων για να αναγνωρίσει πιο σύνθετα και λεπτομερή μοτίβα.

Η διαδικασία ξεκινά με την εφαρμογή ενός φίλτρου που έχει σχεδιαστεί για να ανιχνεύει συγκεκριμένα χαρακτηριστικά πάνω στην εικόνα εισόδου, σε μια διαδικασία που ονομάζεται συνέλιξη. Το αποτέλεσμα αυτής της διαδικασίας είναι ένας χάρτης(map) χαρακτηριστικών, ο οποίος επισημαίνει την παρουσία των εντοπισμένων χαρακτηριστικών στην εικόνα. Αυτός ο χάρτης χαρακτηριστικών χρησιμοποιείται στη συνέχεια ως είσοδος για το επόμενο επίπεδο, επιτρέποντας στο CNN να δημιουργεί σταδιακά μια ιεραρχική αναπαράσταση της εικόνας.

3.4.2 Αρχιτεκτονική του CNN

Ένα CNN συνήθως περιλαμβάνει διάφορα επίπεδα, τα οποία χωρίζονται γενικά σε τρεις κατηγορίες:

- 1) επίπεδα συνελίξεων (convolutional layers),
- 2) επίπεδα συγκέντρωσης (pooling layers) και
- 3) πλήρως συνδεδεμένα επίπεδα (fully connected layers)

Καθώς τα δεδομένα περνούν από αυτά τα στρώματα, η πολυπλοκότητα του CNN αυξάνεται, επιτρέποντάς του να αναγνωρίζει σταδιακά μεγαλύτερα τμήματα μιας εικόνας και πιο αφηρημένα χαρακτηριστικά.

Convolutional layer

Το στρώμα συνελίξεων αποτελεί το βασικό στοιχείο ενός CNN και είναι υπεύθυνο για την εκτέλεση των περισσότερων υπολογισμών. Αυτό το στρώμα χρησιμοποιεί ένα φίλτρο ή πυρήνα, που είναι ένας μικρός πίνακας βαρών, το οποίο κινείται στο δεκτικό πεδίο της εικόνας εισόδου για να ανιχνεύσει συγκεκριμένα χαρακτηριστικά.

Η διαδικασία ξεκινά με την κίνηση του πυρήνα κατά μήκος του πλάτους και του ύψους της εικόνας, σαρώνοντας τελικά ολόκληρη την εικόνα με πολλαπλές επαναλήψεις. Σε κάθε θέση, υπολογίζεται το γινόμενο κουκκίδων μεταξύ των βαρών του πυρήνα και των τιμών των εικονοστοιχείων της εικόνας κάτω από τον πυρήνα. Έτσι, η εικόνα εισόδου μετατρέπεται σε ένα σύνολο χαρτών χαρακτηριστικών, όπου κάθε χάρτης αντιπροσωπεύει την παρουσία και την ένταση ενός συγκεκριμένου χαρακτηριστικού σε διάφορα σημεία της εικόνας.

Τα CNN συνήθως περιλαμβάνουν πολλαπλά στοιβαγμένα επίπεδα συνελικτικής ανάλυσης. Μέσα από αυτήν την πολυεπίπεδη αρχιτεκτονική, το CNN αποκωδικοποιεί σταδιακά τις οπτικές πληροφορίες που υπάρχουν στα αρχικά δεδομένα της εικόνας. Στα αρχικά επίπεδα, το CNN εντοπίζει βασικά χαρακτηριστικά, όπως ακμές, υφές και χρώματα. Στα βαθύτερα επίπεδα, τα δεδομένα που προέρχονται από τους χάρτες χαρακτηριστικών των προηγούμενων επιπέδων επιτρέπουν στο CNN να ανιχνεύει πιο σύνθετα μοτίβα, αντικείμενα και σκηνές.

Pooling layer

Το στρώμα συγκέντρωσης σε ένα CNN είναι ένα σημαντικό στοιχείο που ακολουθεί το στρώμα συνελίξεων. Αν και, όπως και το στρώμα συνελίξεων, σαρώνει την εικόνα εισόδου, η λειτουργία του είναι διαφορετική. Αντί να εντοπίζει χαρακτηριστικά, το στρώμα συγκέντρωσης συμπυκνώνει τις πληροφορίες, μειώνοντας τη διάσταση των δεδομένων και διατηρώντας τα πιο σημαντικά χαρακτηριστικά, κάτι που βοηθά στη μείωση της πολυπλοκότητας και του υπολογιστικού κόστους.

Το στρώμα συγκέντρωσης έχει ως στόχο να μειώσει τη διάσταση των δεδομένων εισόδου, διατηρώντας ωστόσο τις σημαντικές πληροφορίες, γεγονός που ενισχύει την απόδοση του

δικτύου. Αυτό επιτυγχάνεται συνήθως με τη διαδικασία της μείωσης δειγματοληψίας, που περιλαμβάνει τη μείωση του αριθμού των σημείων δεδομένων στην είσοδο.

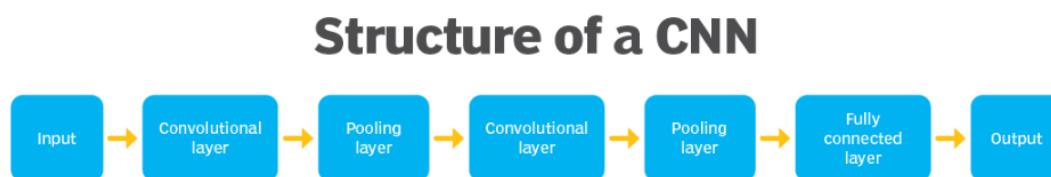
Για τα CNNs, η μείωση της διάστασης συνήθως σημαίνει μείωση του αριθμού των εικονοστοιχείων που χρησιμοποιούνται για την αναπαράσταση της εικόνας. Η πιο συνηθισμένη μέθοδος συγκέντρωσης είναι το max pooling, το οποίο διατηρεί τη μέγιστη τιμή μέσα σε ένα καθορισμένο παράθυρο (δηλαδή το μέγεθος του πυρήνα) και απορρίπτει τις υπόλοιπες τιμές. Μια άλλη κοινή τεχνική είναι το average pooling, που ακολουθεί παρόμοια διαδικασία αλλά χρησιμοποιεί τη μέση τιμή αντί της μέγιστης τιμής.

Fully connected layer

Το πλήρως συνδεδεμένο στρώμα παίζει κρίσιμο ρόλο στα τελικά στάδια ενός CNN, καθώς είναι υπεύθυνο για την ταξινόμηση των εικόνων με βάση τα χαρακτηριστικά που έχουν εξαχθεί από τα προηγούμενα στρώματα. Ο όρος «πλήρως συνδεδεμένος» υποδηλώνει ότι κάθε νευρώνας σε αυτό το στρώμα συνδέεται με όλους τους νευρώνες στο επόμενο στρώμα.

Το πλήρως συνδεδεμένο στρώμα συνδυάζει τα διάφορα χαρακτηριστικά που έχουν εξαχθεί από τα προηγούμενα στρώματα συνελίξεων(convolutional layer) και ομαδοποίησης(pooling layer), και τα αντιστοιχίζει σε συγκεκριμένες κλάσεις ή αποτελέσματα. Κάθε είσοδος από το προηγούμενο στρώμα συνδέεται με όλες τις μονάδες ενεργοποίησης στο πλήρως συνδεδεμένο στρώμα, επιτρέποντας στο CNN να αξιολογεί συνολικά όλα τα χαρακτηριστικά προτού πάρει την τελική απόφαση ταξινόμησης.

Παρακάτω παρουσιάζεται και μία ενδεικτική απεικόνιση της συγκεκριμένης αρχιτεκτονικής ενός CNN.



Σχήμα 3.1 Δομή ενός CNN[7]

Αξίζει να τονιστεί ότι, τα CNNs διαφέρουν από τα παραδοσιακά νευρωνικά δίκτυα(NNs) σε ορισμένα βασικά σημεία. Ειδικότερα, σε ένα CNN, δεν υπάρχει σύνδεση κάθε κόμβου σε ένα επίπεδο με κάθε κόμβο στο επόμενο επίπεδο. Επειδή τα συνελκτικά στρώματα έχουν λιγότερες παραμέτρους σε σύγκριση με τα πλήρως συνδεδεμένα στρώματα ενός παραδοσιακού νευρωνικού δικτύου, τα CNN είναι πιο αποτελεσματικά στην επεξεργασία εικόνας.

Τα CNN χρησιμοποιούν μια τεχνική γνωστή ως διαμοιρασμός παραμέτρων(parameter sharing), η οποία τα καθιστά πολύ πιο αποδοτικά στην επεξεργασία δεδομένων εικόνας. Στα στρώματα συνελκτικής σύνδεσης, το ίδιο φίλτρο, με σταθερά βάρη, χρησιμοποιείται για τη σάρωση ολόκληρης της εικόνας, γεγονός που μειώνει σημαντικά τον αριθμό των παραμέτρων σε σύγκριση με ένα πλήρως συνδεδεμένο στρώμα ενός παραδοσιακού νευρωνικού δικτύου. Επιπλέον, τα στρώματα συνένωσης μειώνουν περαιτέρω τη διάσταση των δεδομένων, βελτιώνοντας έτσι την συνολική αποδοτικότητα και την ικανότητα γενίκευσης του CNN.

3.4.3 Πεδία εφαρμογής των CNNs

Δεδομένου ότι η επεξεργασία οπτικών δεδομένων αποτελεί μία συχνή ρουτίνα, τα CNNs έχουν ευρύ φάσμα εφαρμογών στον πραγματικό κόσμο.

Πιο συγκεκριμένα χρησιμοποιούνται σε τομείς όπως :

Υγειονομική περίθαλψη: Στον τομέα της υγειονομικής περίθαλψης, τα CNNs χρησιμοποιούνται για να υποστηρίξουν τη διάγνωση και την απεικόνιση ιατρικών εικόνων. Για παράδειγμα, ένα CNN μπορεί να αναλύσει ακτινογραφίες για να ανιχνεύσει ανωμαλίες που ενδέχεται να υποδεικνύουν ασθένειες, βοηθώντας έτσι στη διάγνωση και το σχεδιασμό της θεραπείας.

Αυτοκίνηση : Στην αυτοκινητοβιομηχανία, τα CNNs εφαρμόζονται σε οχήματα για να ερμηνεύσουν δεδομένα από κάμερες και αισθητήρες, βοηθώντας τα οχήματα να πλοηγηθούν στο περιβάλλον τους. Επίσης, χρησιμοποιούνται σε ημι-αυτόνομα οχήματα για χαρακτηριστικά όπως ο αυτοματοποιημένος έλεγχος ταχύτητας και η βοήθεια στάθμευσης.

Social media : Στις πλατφόρμες κοινωνικής δικτύωσης, τα CNNs χρησιμοποιούνται για διάφορες μεθόδους ανάλυσης εικόνας. Για παράδειγμα, μπορούν να προτείνουν άτομα που πρέπει να αναγνωριστούν σε φωτογραφίες.

Λιανικό εμπόριο: Στον τομέα του λιανικού εμπορίου, οι εταιρείες χρησιμοποιούν CNNs για συστήματα οπτικής αναζήτησης, που επιτρέπουν στους χρήστες να αναζητούν προϊόντα με

εικόνας αντί για κείμενο. Επίσης, βοηθούν στη βελτίωση των συστημάτων συστάσεων, εντοπίζοντας προϊόντα που μοιάζουν οπτικά με εκείνα που έχει ενδιαφέρον ο αγοραστής.

3.5.1 Αναδρομικά Νευρωνικά Δίκτυα (Recurrent Neural Networks – RNNs)

Το αναδρομικό νευρωνικό δίκτυο είναι μια μορφή τεχνητού νευρωνικού δικτύου που χρησιμοποιείται κυρίως στην αναγνώριση ομιλίας και στην επεξεργασία φυσικής γλώσσας. Αυτά τα δίκτυα είναι σε θέση να κατανοούν την αλληλουχία των δεδομένων και να αξιοποιούν τα πρότυπα που αναγνωρίζουν για να προβλέψουν το επόμενο πιθανό συμβάν.

Τα RNNs αξιοποιούνται στη βαθιά μάθηση(DL) και στη δημιουργία μοντέλων. Είναι ιδιαίτερα αποτελεσματικά σε εφαρμογές όπου το πλαίσιο(context) παίζει καθοριστικό ρόλο για την πρόβλεψη ενός αποτελέσματος. Σε αντίθεση με άλλους τύπους τεχνητών νευρωνικών δικτύων, τα RNNs χρησιμοποιούν βρόχους ανατροφοδότησης για να επεξεργαστούν ακολουθίες δεδομένων, κάτι που επηρεάζει την τελική έξοδο. Αυτοί οι βρόχοι ανατροφοδότησης επιτρέπουν τη διατήρηση πληροφοριών, μια διαδικασία που συχνά αναφέρεται ως μνήμη.

Οι εφαρμογές των RNNs συνδέονται κυρίως με γλωσσικά μοντέλα, όπου η πρόβλεψη του επόμενου γράμματος σε μια λέξη ή της επόμενης λέξης σε μια πρόταση βασίζεται στις προηγούμενες πληροφορίες. Η γραφή που παράγεται από τα RNN αποτελεί μια μορφή υπολογιστικής δημιουργικότητας. Αυτή η παραγόμενη γραφή ανταπεξέρχεται αρκετά με την ανθρώπινη δημιουργικότητα καθώς το δίκτυο μπορεί να κατανοεί τη γραμματική και τη σημασία, όπως τις έχει μάθει από το σύνολο εκπαίδευσής του.

Τρόπος Εκπαίδευσης των RNNs

Τα τεχνητά νευρωνικά δίκτυα αποτελούνται από διασυνδεδεμένα στοιχεία επεξεργασίας δεδομένων, τα οποία είναι σχεδιασμένα με τρόπο που να προσομοιώνει τη λειτουργία του ανθρώπινου εγκεφάλου. Αυτά τα δίκτυα περιλαμβάνουν στρώματα τεχνητών νευρώνων, δηλαδή κόμβων, που μπορούν να επεξεργάζονται τις εισόδους και να μεταδίδουν τις εξόδους τους σε άλλους κόμβους. Οι κόμβοι αυτοί συνδέονται μέσω ακμών, τα οποία καθορίζουν την ένταση του σήματος και επηρεάζουν την τελική έξοδο του δικτύου.

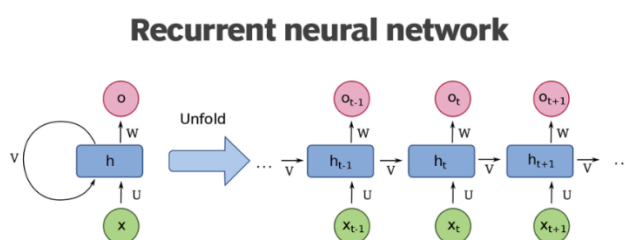
Σε ορισμένες περιπτώσεις, τα τεχνητά νευρωνικά δίκτυα επεξεργάζονται δεδομένα μόνο σε μία κατεύθυνση, από την είσοδο προς την έξοδο. Αυτά τα δίκτυα, γνωστά ως «νευρωνικά δίκτυα τροφοδότησης προς τα εμπρός» (feed-forward), περιλαμβάνουν και τα συνελκτικά νευρωνικά δίκτυα, τα οποία χρησιμοποιούνται κυρίως σε συστήματα αναγνώρισης εικόνων. Αντίθετα, τα RNN έχουν τη δυνατότητα να διαστρώνονται ώστε να επεξεργάζονται πληροφορίες προς δύο κατευθύνσεις.

Όπως και τα νευρωνικά δίκτυα <<feed-forward>>, τα RNNs μπορούν να επεξεργάζονται δεδομένα από την αρχική είσοδο μέχρι την τελική έξοδο. Ωστόσο, σε αντίθεση με τα δίκτυα τροφοδότησης προς τα εμπρός, τα RNNs χρησιμοποιούν βρόχους ανατροφοδότησης, όπως η οπισθοδιάδοση(backpropagation) μέσω του χρόνου, κατά τη διάρκεια της υπολογιστικής διαδικασίας, για να επαναφέρουν πληροφορίες στο δίκτυο. Αυτοί οι βρόχοι συνδέουν τις

εισόδους, επιτρέποντας στα RNNs να επεξεργάζονται σειριακά και χρονικά δεδομένα με αποτελεσματικότητα.

Ένα νευρωνικό δίκτυο με οπισθοδιάδοση μέσω του χρόνου είναι ένα RNN όπου ο αριθμός των χρονικών βημάτων στην ακολουθία εισόδου περιορίζεται με τη μείωση της ακολουθίας εισόδου. Αυτό είναι ιδιαίτερα χρήσιμο σε αναδρομικά νευρωνικά δίκτυα που λειτουργούν ως μοντέλα ακολουθίας προς ακολουθία, όπου η ακολουθία εισόδου έχει περισσότερα χρονικά βήματα από την ακολουθία εξόδου. Αυτή η προσέγγιση επιτρέπει τη διαχείριση μακρών ακολουθιών, διατηρώντας την αποτελεσματικότητα του μοντέλου.

Παρακάτω παρουσιάζεται ένα διάγραμμα μιας μονάδας RNN όπου το κάτω μέρος είναι η είσοδος(input), το μεσαίο μέρος η κρυφή κατάσταση(hidden state) και το πάνω μέρος η έξοδος(output). Τα σύμβολα U, V, W αποτελούν τα βάρη του δικτύου.



Σχήμα 3.2 Διάγραμμα μιας μονάδας RNN[7]

3.5.2 Παραγωγικό Αντιφατικό Δίκτυο (Generative Adversarial Networks – GANs)

Ένα παραγωγικό αντιφατικό δίκτυο (GAN) είναι ένα μοντέλο μηχανικής μάθησης (ML) όπου δύο νευρωνικά δίκτυα συναγωνίζονται το ένα το άλλο, αξιοποιώντας τεχνικές βαθιάς μάθησης για να βελτιώσουν την ακρίβεια των προβλέψεών τους. Τα GANs συνήθως λειτουργούν χωρίς επίβλεψη και βασίζονται σε ένα ανταγωνιστικό πλαίσιο μηδενικού αθροίσματος, όπου το κέρδος του ενός δικτύου αντιστοιχεί στη ζημία του άλλου.

Τα δύο νευρωνικά δίκτυα που αποτελούν ένα GAN ονομάζονται γεννήτρια(generator) και διαχωριστής(discriminator). Η γεννήτρια είναι ένα συνελκτικό νευρωνικό δίκτυο, ενώ ο διαχωριστής είναι ένα αποσυμβολικό(deconvolutional) νευρωνικό δίκτυο. Η γεννήτρια έχει ως στόχο να δημιουργεί τεχνητά δεδομένα που μοιάζουν όσο το δυνατόν περισσότερο με πραγματικά. Από την άλλη πλευρά, ο στόχος του διαχωριστή είναι να αναγνωρίζει ποιες από τις εισόδους που λαμβάνει είναι τεχνητά κατασκευασμένες.

Πιο συγκεκριμένα, τα παραγωγικά μοντέλα δημιουργούν τα δικά τους δεδομένα για την εκπαίδευσή τους. Η γεννήτρια μαθαίνει να παράγει ψευδή δεδομένα, ενώ το δίκτυο διάκρισης εκπαιδεύεται να ξεχωρίζει τα ψευδή δεδομένα της γεννήτριας από τα αληθινά παραδείγματα. Όταν ο διαχωριστής αναγνωρίζει γρήγορα τα ψευδή δεδομένα που δημιουργεί η γεννήτρια, όπως μια εικόνα που δεν αντιστοιχεί σε ανθρώπινο πρόσωπο, η γεννήτρια τιμωρείται. Με την επαναλαμβανόμενη διαδικασία ανατροφοδότησης μεταξύ των δύο δικτύων, η γεννήτρια αρχίζει να παράγει πιο ποιοτικά και αληθοφανή αποτελέσματα, ενώ ο διαχωριστής βελτιώνεται στην ανίχνευση των τεχνητών δεδομένων.

3.5.3 Λειτουργία των GANs

Τα GANs μπορούν να χωριστούν σε τρεις κύριες κατηγορίες:

- 1) Γενετικά:** Αυτή η κατηγορία αφορά τον τρόπο με τον οποίο παράγονται τα δεδομένα με βάση ένα πιθανοτικό μοντέλο.
- 2) Αντιθετικά:** Αυτή η κατηγορία περιγράφει τη διαδικασία εκπαίδευσης ενός μοντέλου σε ένα ανταγωνιστικό περιβάλλον.
- 3) Δίκτυα:** Σε αυτή την κατηγορία, χρησιμοποιούνται βαθιά νευρωνικά δίκτυα ως αλγόριθμοι τεχνητής νοημοσύνης για σκοπούς εκπαίδευσης.

Το πρώτο βήμα για τη δημιουργία ενός GAN είναι να οριστεί η επιθυμητή τελική έξοδος και να συλλεχθεί ένα αρχικό σύνολο δεδομένων εκπαίδευσης που να ανταποκρίνεται σε αυτές τις προδιαγραφές. Στη συνέχεια, τα δεδομένα αυτά λαμβάνουν μια τυχαία διάταξη και εισάγονται στη γεννήτρια, η οποία εκπαιδεύεται μέχρι να επιτύχει μια βασική ακρίβεια στην παραγωγή των επιθυμητών εξόδων.

Στη συνέχεια, τα δείγματα ή οι εικόνες που παράγονται από τη γεννήτρια τροφοδοτούνται στον διαχωριστή μαζί με πραγματικά δεδομένα από το αρχικό σύνολο. Μετά την επεξεργασία αυτών των δεδομένων από τα μοντέλα της γεννήτριας και του διαχωριστή, ξεκινά η διαδικασία βελτιστοποίησης μέσω backpropagation. Ο διαχωριστής αξιολογεί τις εικόνες και επιστρέφει μια πιθανότητα μεταξύ 0 και 1 για να δείξει την αυθεντικότητα κάθε εικόνας, όπου το 1 αντιπροσωπεύει τις πραγματικές εικόνες και το 0 τις ψεύτικες. Στη συνέχεια, οι τιμές αυτές αξιολογούνται χειροκίνητα και η διαδικασία επαναλαμβάνεται μέχρι να επιτευχθεί το επιθυμητό αποτέλεσμα.

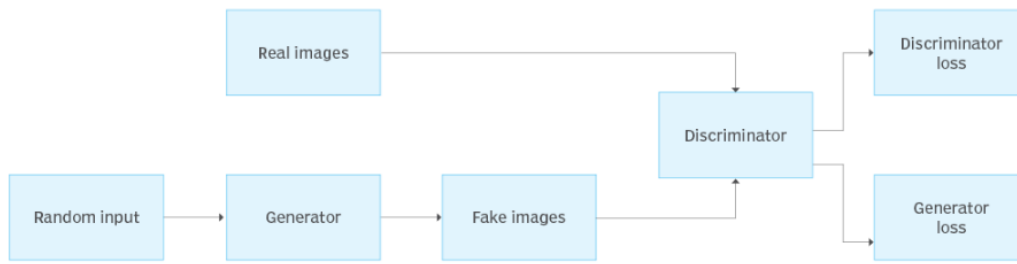
Συνοψίζοντας ένα GAN ακολουθεί τα εξής βήματα:

- 1) Η γεννήτρια δημιουργεί μια εικόνα χρησιμοποιώντας τυχαίους αριθμούς ως είσοδο.**
- 2) Ο διαχωριστής λαμβάνει την εικόνα που παρήγαγε η γεννήτρια, μαζί με μια ροή φωτογραφιών από το πραγματικό σύνολο δεδομένων.**
- 3) Ο διαχωριστής επεξεργάζεται τόσο τις πραγματικές όσο και τις ψεύτικες εικόνες και παράγει πιθανότητες, με τιμές μεταξύ 0 και 1, όπου το 1 σηματοδοτεί την αυθεντικότητα και το 0 την απομίμηση.**

Αυτό δημιουργεί έναν διπλό βρόχο ανατροφοδότησης: ο διαχωριστής ανατροφοδοτείται από την πραγματική βάση των εικόνων, ενώ η γεννήτρια ανατροφοδοτείται από τον διαχωριστή.

Παρακάτω απεικονίζεται η ανάλογη δομή ενός GAN :

The structure of a GAN



Σχήμα 3.3 Δομή ενός GAN[7]

3.5.4 Κατηγορίες GANs

Τα GAN υπάρχουν σε πολλές μορφές και μπορούν να αξιοποιηθούν σε διάφορες διαδικασίες. Οι πιο δημοφιλείς τύποι GAN είναι :

Vanilla GAN: Πρόκειται για την πιο απλή μορφή όλων των GAN. Ο αλγόριθμός του στοχεύει στη βελτιστοποίηση μιας μαθηματικής εξίσωσης μέσω της στοχαστικής καθόδου κλίσης, μιας μεθόδου μάθησης που επεξεργάζεται το σύνολο των δεδομένων περνώντας από ένα παράδειγμα κάθε φορά. Αποτελείται από δύο βασικά μέρη: μια γεννήτρια και έναν διαχωριστή. Η γεννήτρια και ο διαχωριστής είναι απλά πολυεπίπεδα perceptrons (νευρωνικά δίκτυα), με τη γεννήτρια να μαθαίνει την κατανομή των δεδομένων και να δημιουργεί νέες εικόνες, ενώ ο διαχωριστής προσπαθεί να εκτιμήσει την πιθανότητα η είσοδος να ανήκει σε μια συγκεκριμένη κλάση, δηλαδή να προσδιορίσει αν είναι πραγματική ή ψεύτικη.

Conditional GAN (cGAN): Αυτός ο τύπος GAN ενσωματώνει ετικέτες τάξης για να προσφέρει στο δίκτυο νέες και συγκεκριμένες πληροφορίες. Κατά την εκπαίδευση του cGAN, το δίκτυο δέχεται εικόνες με τις αντίστοιχες ετικέτες τους. Αυτό επιτρέπει στη γεννήτρια να δημιουργεί εικόνες που αντιστοιχούν σε συγκεκριμένες κατηγορίες και στον διαχωριστή να μάθει πώς να διακρίνει μεταξύ των διαφορετικών ετικετών με μεγαλύτερη ακρίβεια.

Deep Convolutional GAN (DCGAN): Αυτός ο τύπος GAN χρησιμοποιεί ένα βαθύ συνελκτικό νευρωνικό δίκτυο για την παραγωγή εικόνων υψηλής ανάλυσης με καλή διαφοροποίηση. Οι συνελίξεις, μια τεχνική που αντλεί σημαντικές πληροφορίες από τα δεδομένα, λειτουργούν ιδιαίτερα καλά με τις εικόνες, επιτρέποντας στο δίκτυο να απορροφήσει γρήγορα τις βασικές λεπτομέρειες και να δημιουργήσει πιο ποιοτικά και ρεαλιστικά αποτελέσματα.

CycleGAN: Αυτή η αρχιτεκτονική GAN είναι ευρέως χρησιμοποιούμενη για την εκμάθηση μετασχηματισμών μεταξύ εικόνων διαφορετικών στυλ. Για παράδειγμα, μπορεί να μάθει πώς να μετατρέπει μια εικόνα από το χειμώνα σε καλοκαίρι.

StyleGAN: Το StyleGAN εισάγει σημαντικές βελτιώσεις στις αρχικές αρχιτεκτονικές μοντέλων γεννήτριας. Το StyleGAN είναι ικανό να δημιουργεί φωτορεαλιστικές, υψηλής ποιότητας εικόνες προσώπων. Επιπλέον, επιτρέπει στους χρήστες να τροποποιούν το μοντέλο για να προσαρμόσουν την εμφάνιση των παραγόμενων εικόνων, προσφέροντας έτσι μεγαλύτερη ευελιξία και έλεγχο στην παραγωγή των αποτελεσμάτων.

3.5.4 Πεδία εφαρμογής των GANs

Τα GANs μπορούν να χρησιμοποιηθούν για διάφορες διαδικασίες, όπως:

- 1) Ανίχνευση ανωμαλιών:** Εντοπισμός και ανάλυση ασυνήθιστων ή μη φυσιολογικών μοτίβων σε δεδομένα.
- 2) Επαύξηση δεδομένων:** Δημιουργία πρόσθετων δεδομένων για την ενίσχυση της ποικιλίας και ποσότητας των εκπαιδευτικών συνόλων.
- 3) Σύνθεση εικόνων:** Δημιουργία νέων εικόνων συνδυάζοντας ή τροποποιώντας υπάρχουσες εικόνες.
- 4) Μετάφραση κειμένου σε εικόνα:** Δημιουργία εικόνων που απεικονίζουν σκηνές ή αντικείμενα με βάση περιγραφές κειμένου.
- 5) Μετάφραση εικόνας σε εικόνα:** Μετατροπή μιας εικόνας σε άλλη μορφή ή στυλ, όπως η αλλαγή στυλ ή η ενίσχυση της ανάλυσης.

Οι συνηθισμένες περιπτώσεις χρήσης των GAN περιλαμβάνουν:

- 1) Συμπλήρωση εικόνων από ένα περίγραμμα
- 2) Δημιουργία ρεαλιστικών εικόνων από κείμενο
- 3) Παραγωγή φωτορεαλιστικών απεικονίσεων πρωτοτύπων προϊόντων
- 4) Μετατροπή ασπρόμαυρων εικόνων σε έγχρωμες
- 5) Μεταφράσεις φωτογραφιών από σκίτσα ή σημασιολογικές εικόνες
- 6) Πρόβλεψη επακόλουθων καρτέ βίντεο: Εκτίμηση και δημιουργία των επόμενων καρτέ σε μια ακολουθία βίντεο.
- 7) Δημιουργία βαθιάς απομίμησης(deepfake): Παραγωγή βίντεο που αναπαράγουν ή μιμούνται συγκεκριμένα στυλ ή χαρακτηριστικά.

Κεφάλαιο 4 : Οπτικά δίκτυα για ΑΙ εφαρμογές

4.1 Εισαγωγή στα φωτονικά ολοκληρωμένα και ενεργειακά αποδοτικά δίκτυα στα κέντρα δεδομένων(PINE)

Η ραγδαία ανάπτυξη των εφαρμογών ανάλυσης δεδομένων που αξιοποιούν τεχνικές μηχανικής μάθησης οδηγεί στη σύγκλιση μεταξύ κέντρων δεδομένων και συστημάτων υπολογιστών υψηλής απόδοσης (HPC). Αυτό αυξάνει σημαντικά την ανάγκη για πιο αποδοτικούς υπολογιστικούς πόρους. Η αποδοτική εκτέλεση αυτών των παράλληλων αρχιτεκτονικών εξαρτάται σε μεγάλο βαθμό από τον τρόπο μεταφοράς των δεδομένων μεταξύ των υπολογιστών, της μνήμης και των αποθηκευτικών πόρων, και επηρεάζεται αρνητικά από την αυξανόμενη κατανάλωση ενέργειας που απαιτείται για τη μετακίνηση των δεδομένων, όπως επίσης και από την αποδοτικότητα στη χρήση των ετερογενών υπολογιστικών και μνημονικών πόρων.

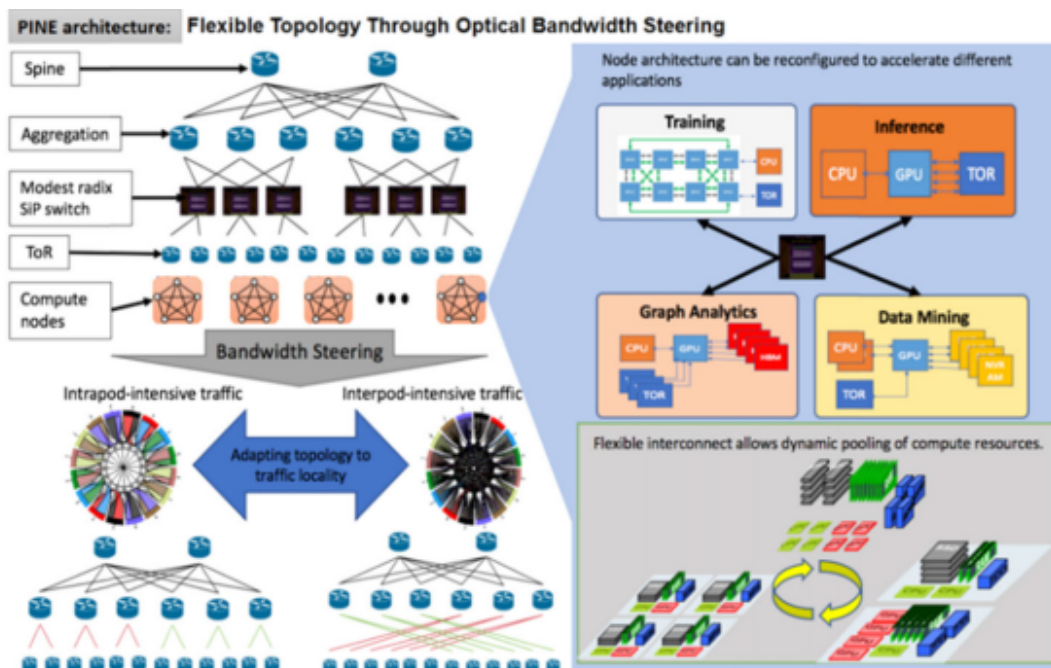
Αυτές οι προκλήσεις ώθησαν την έρευνα στην κατεύθυνση των φωτονικών ολοκληρωμένων και ενεργειακά αποδοτικών δικτύων στα κέντρα δεδομένων (Photonic Integrated Networked Energy Efficient Datacenter - PINE). Αυτό το ερευνητικό πρόγραμμα αποτελεί μέρος του προγράμματος ARPA-E, με το όνομα ENLITENED (Energy-efficient Light-wave Integrated Technology Enabling Networks that Enhance Dataprocessing). Η αρχιτεκτονική PINE προσπαθεί να λύσει το πρόβλημα της μεταφοράς δεδομένων, αξιοποιώντας τις μοναδικές δυνατότητες της φωτονικής, ώστε να κατευθύνει το εύρος ζώνης μόνο εκεί που είναι απαραίτητο, αντί να υπερφορτώνει το δίκτυο, κάτι που αυξάνει δραματικά την κατανάλωση ενέργειας. Επίσης, η φωτονική μπορεί να συμβάλει στην αποδοτική διαχείριση των πόρων. Το PINE στηρίζεται σε τρεις βασικούς άξονες:

- 1) Η αποσυνδυστασμένη αρχιτεκτονική PINE προτείνει μια νέα προσέγγιση για τα κέντρα δεδομένων, με στόχο τη μεγιστοποίηση των πλεονεκτημάτων που προκύπτουν από την ομαλή ενσωμάτωση χαμηλής ισχύος φωτονικών συνδέσεων πυριτίου (SiP) και την χρήση μεγάλου αριθμού ευρυζωνικών φωτονικών μεταγωγέων με χαμηλή ακτίνα. Αυτή η αρχιτεκτονική επιτρέπει τη διασύνδεση πολλαπλών τσιπ (MCM) και την αναδιαμόρφωσή τους εντός του ίδιου rack, επιτυγχάνοντας υψηλή απόδοση με χαμηλό κόστος καταμερισμού πόρων σε επίπεδο rack. Η κατανομή των πόρων στην αρχιτεκτονική PINE προβάλλει το κέντρο δεδομένων ως μια δεξαμενή διαχωρισμένων πόρων, οι οποίοι μπορούν να ανακατανεμηθούν με λεπτομερή ευελιξία, αποφεύγοντας τη συμφόρηση των εφαρμογών σε έναν συγκεκριμένο τύπο πόρου και ταυτόχρονα προλαμβάνοντας την υποαξιοποίηση των διαθέσιμων πόρων.
- 2) Τα φωτονικά MCMs με εξαιρετικά χαμηλή κατανάλωση ενέργειας και υψηλή πυκνότητα εύρους ζώνης προσφέρουν λύση στο πρόβλημα του "chip escape". Οι σχεδιαστές συστημάτων συχνά βρίσκονται περιορισμένοι από τα όρια της μνήμης και της συσκευασίας εισόδου/εξόδου (I/O). Η αύξηση του εύρους ζώνης I/O συνοδεύεται

από υψηλό ενεργειακό κόστος. Παρόλο που πολλοί πυρήνες CPU/GPU έχουν την ικανότητα να εκτελούν εξαιρετικά απαιτητικές υπολογιστικές διεργασίες, συχνά δεν διαθέτουν το αναγκαίο εύρος ζώνης εκτός του τσιπ για να αξιοποιήσουν πλήρως τις δυνατότητές τους. Το ενσωματωμένο σύστημα PINE, με την ευέλικτη φωτονική συνδεσιμότητα υψηλής πυκνότητας και εύρους ζώνης, υλοποιημένο μέσω μιας πλατφόρμας ενεργού interposer MCM, επιτρέπει τη διαφυγή δεδομένων από το τσιπ με ρυθμούς που φτάνουν τα πολλαπλά Tbps (terabits per second).

- 3) Η ενεργειακά βελτιστοποιημένη πυκνή φωτονική συνδεσιμότητα μέσω πολυπλεξίας με διαίρεση μήκους κύματος (DWDM) στην τεχνολογία PINE αξιοποιεί μια νέα γενιά εξαρτημάτων που έχουν σχεδιαστεί ειδικά για ενεργειακή αποδοτικότητα. Η πλατφόρμα αυτή περιλαμβάνει μια προηγμένη φωτονική πηγή, η οποία αποτελείται από ένα μοναδικό λέιζερ υψηλής ισχύος και αποδοτικότητας. Το λέιζερ αυτό είναι συνδεδεμένο με ένα σύστημα πολλαπλών μηκών κύματος, επιτρέποντας την κατανομή της ισχύος σε περισσότερα από 50 διαφορετικά μήκη κύματος, προσφέροντας σημαντική ενεργειακή οικονομία και βελτιωμένη απόδοση στη μεταφορά δεδομένων.

Παρακάτω στο Σχήμα 4.1 παρουσιάζεται μία απεικόνιση της αρχιτεκτονικής ενός PINE συστήματος:



Σχήμα 4.1 Αρχιτεκτονική ενός PINE συστήματος[5]

4.2 Αρχιτεκτονική ενός συστήματος PINE

Φάση 1 : Αρχιτεκτονικός σχεδιασμός και αξιολόγηση ενός PINE

Η αρχιτεκτονική του συστήματος PINE επικεντρώθηκε στη διαχείριση του εύρους ζώνης σε όλο το σύστημα, χρησιμοποιώντας μια κλιμακούμενη και ευέλικτη αρχιτεκτονική fat tree. Αυτή η αρχιτεκτονική πέτυχε μέση βελτίωση 2,2 φορές στις συναλλαγές ανά joule, μειώνοντας την κατανάλωση ενέργειας κατά 55% για συναλλαγές σταθερού μεγέθους. Επιπλέον, σημειώθηκε μείωση κατά 20% στον μέσο χρόνο καθυστέρησης δικτύου σε ένα ευρύ φάσμα εφαρμογών .

Τα βασικά στοιχεία της διαχείρισης εύρους ζώνης στο PINE είναι οι επαναδιαμορφώσιμοι switch πυριτίου (SiP). Οι διαδικασίες συναρμολόγησης PINE MCM 2.5D και 3D υλοποιούνται σε πλατφόρμες interposer και 3D ενεργούς interposer, όπου επιτεύχθηκε υψηλής πυκνότητας ηλεκτρονική και φωτονική ολοκλήρωση. Επίσης, αναπτύχθηκε η πρώτη πλατφόρμα φωτονικών διακοπών πολλαπλών επιπέδων, η οποία κατάφερε να πετύχει ρεκόρ χαμηλής διασταυρούμενης ομιλίας και εξάλειψης άνω των 50 dB .

Το πλήρως ενσωματωμένο λέιζερ comb του PINE, με αποδεδειγμένη απόδοση μετατροπής αντλίας ισχύος στο 41%, αποτελεί το βασικό στοιχείο που υποστηρίζει τις συνδέσεις στη Φάση 1 του PINE. Η αρχιτεκτονική του PINE βελτιώνει περαιτέρω την επεκτασιμότητα με τη χρήση οπτικών ενισχυτών ημιαγωγών (SOA) υψηλής απόδοσης, καθώς και του πρώτου μονολιθικού SOA κβαντικής τελείας (QD) σε πυρίτιο, που πέτυχε ρεκόρ απόδοσης wall-plug (WPE) στο 14,2% και κέρδος στο τσιπ 39 dB.

Επιπλέον, η πλατφόρμα MCM προχώρησε σε προηγμένη τεχνολογία παθητικής ευθυγράμμισης υψηλής πυκνότητας για την οπτική σύνδεση chip-I/O με οπτικές ίνες. Αυτή η τεχνολογία έχει αποδειχθεί ανθεκτική τόσο σε μεταβολές θερμοκρασίας όσο και σε κατασκευαστικές παραλλαγές, διατηρώντας ποιινή απόδοσης μικρότερη από 0,6 dB στην απόδοση σύζευξης.

Η Φάση 1 περιλαμβάνει μια λεπτομερή διασταυρούμενη μοντελοποίηση και ανάλυση ενέργειας και απόδοσης σε επίπεδα σύνδεσης, κόμβου και ολόκληρου του συστήματος. Στη Φάση 1 του PINE, παρουσιάστηκε μια ενεργειακά βελτιστοποιημένη σχεδίαση ζεύξης SiP υψηλής πυκνότητας και εύρους ζώνης, με συνολικό εύρος ζώνης 640 Gb/s. Για τη διαίρεση των καναλιών πολυπλεξίας με διαίρεση μήκους κύματος (WDM) σε τέσσερις υποομάδες, χρησιμοποιήθηκε ένας αποδιαπλέκτης δύο σταδίων, με κάθε υποομάδα να περιλαμβάνει 10 διαδοχικούς αντηχείς δακτυλίου διαμορφωμένους στα 16 Gb/s ($4 \times 10 \times 16$ Gb/s). Η ζεύξη λειτουργεί με ενεργειακή κατανάλωση 2,2 pJ/bit, χρησιμοποιώντας το PDK της AIM Photonics, ένα ειδικά σχεδιασμένο συζεύκτη και έναν οδηγό/δέκτη ASIC που αναπτύχθηκε από την ομάδα μας. Ο δέκτης της Φάσης 1 λειτουργεί με ευαισθησία -17 dBm και κατανάλωση 0,75 pJ/bit.

Η Φάση 2 του PINE θα βασιστεί σε αυτές τις επιτυχίες για να μειώσει περαιτέρω την κατανάλωση ενέργειας των δεκτών κατά 80%. Ο νέος δέκτης της Φάσης 2 θα έχει βελτιωμένη ευαισθησία -22,5 dBm, μειώνοντας έτσι την απαιτούμενη ισχύ εισόδου του λέιζερ. Μαζί με βελτιωμένους ειδικά σχεδιασμένους θερμαντήρες και οπτικούς ζεύκτες, η ζεύξη της Φάσης 2 θα στοχεύει σε ενέργεια 1 pJ/bit.

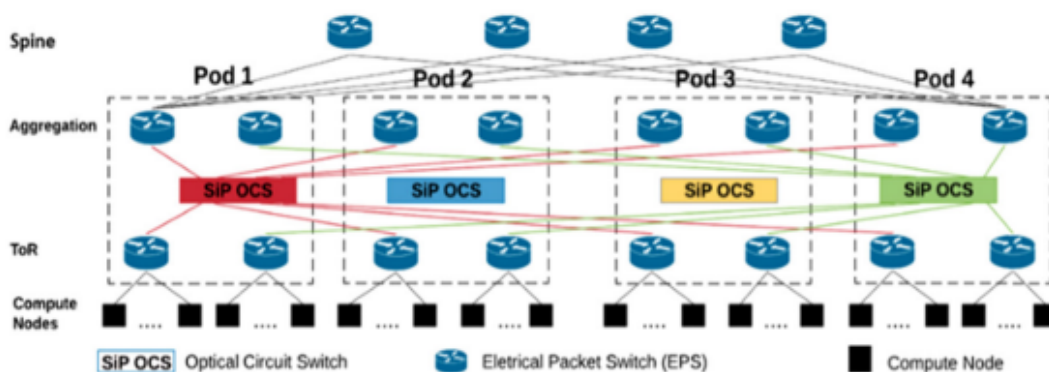
Επιπλέον, η Φάση 2 του PINE θα πραγματοποιήσει ολοκλήρωση σε επίπεδο συστήματος, που θα περιλαμβάνει φωτονικά διασυνδεδεμένα MCM με ευελιξία μεταγωγής, επιδεικνύοντας την αρχιτεκτονική PINE υπό ρεαλιστικούς φόρτους εργασίας. Η Φάση 2 θα επιδιώξει φωτονικές συνδέσεις με συνολικό εύρος ζώνης άνω των 1 Tb/s, θα επιδείξει πυκνότητες εύρους ζώνης MCM στην άκρη του τσιπ άνω των 5 Tb/s/mm και θα μειώσει την κατανάλωση ενέργειας σε ολόκληρο το σύστημα των κέντρων δεδομένων κατά παράγοντα 2,8.

Καθοδήγηση εύρους ζώνης σε όλο το σύστημα PINE

Κατά τη διάρκεια της Φάσης 1, αναπτύξαμε τη στρατηγική καθοδήγησης του εύρους ζώνης σε ολόκληρο το σύστημα, προκειμένου να επιτευχθεί η απρόσκοπτη αναδιαμόρφωση της τοπολογίας διασύνδεσης, ώστε να προσαρμόζεται σε διάφορα πρότυπα επικοινωνίας και φόρτου εργασίας. Εφαρμόζοντας αυτή τη στρατηγική, καταφέραμε να διατηρήσουμε μεγαλύτερο όγκο κίνησης στα χαμηλότερα επίπεδα της τοπολογίας, αποδεικνύοντας ότι το υψηλό εύρος ζώνης μπορεί να μειωθεί σημαντικά στις συνδέσεις χωρίς να επιφέρει ποινές στην απόδοση.

Με την αναδιάρθρωση της συνδεσιμότητας στα χαμηλότερα επίπεδα, περισσότεροι όγκοι κίνησης μπορούν να κατευθύνονται απευθείας προς το μεσαίο επίπεδο (σημείο συγκέντρωσης) χωρίς να απαιτείται η χρήση του ανώτερου επιπέδου της τοπολογίας. Η φυσική τοπολογία της προσέγγισης απεικονίζεται στο Σχήμα 4.2. Η καθοδήγηση του εύρους ζώνης εφαρμόζεται μόνο στις ανόδους των συνδέσεων, καθώς αυτό αρκεί για να διασφαλιστεί ότι η κυκλοφορία που χρησιμοποιεί αυτές τις κατευθυνόμενες συνδέσεις δεν φτάνει στο ανώτερο επίπεδο του δέντρου fat tree.

Στο πειραματικό σύστημα 32 κόμβων με 4×4 μεταγωγείς SiP και το ευέλικτο fat tree PINE, το οποίο θα περιγραφεί αναλυτικά παρακάτω, πραγματοποιούνται πειράματα με τον αλγόριθμο γυροκινητικού τοροειδούς κώδικα HPC (gyrokinetic toroidal code - GTC), καταδεικνύοντας ότι η χρήση διεύθυνσης εύρους ζώνης επέφερε επιτάχυνση 62% στην εκτέλεση της εφαρμογής.



Σχήμα 4.2 Καθοδήγηση εύρους ζώνης στην ευέλικτη αρχιτεκτονική fat tree PINE[5]

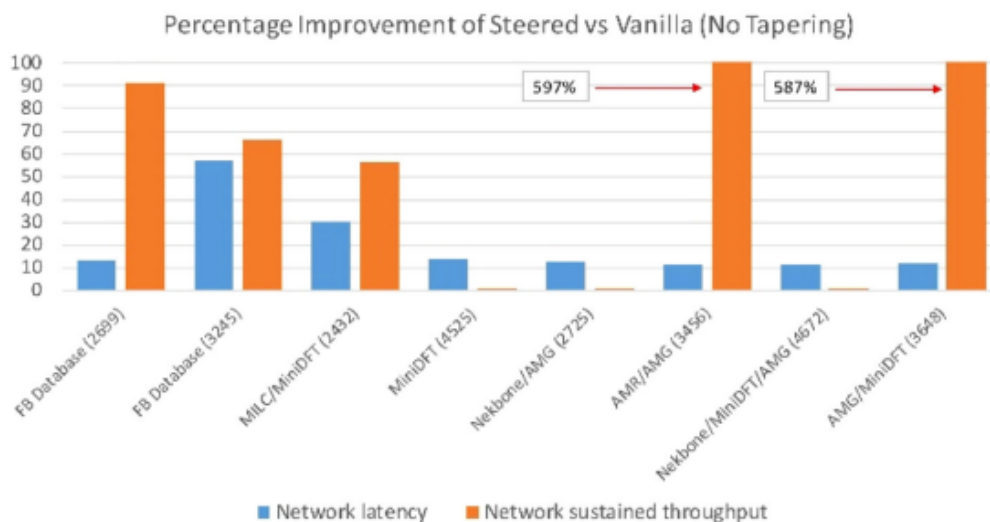
Οι προσομοιώσεις σε επίπεδο συστήματος για την κλιμακούμενη διεύθυνση εύρους ζώνης στην αρχιτεκτονική PINE διεξάγονται χρησιμοποιώντας τον προσομοιωτή δικτύου κύκλων Booksim, ο οποίος έχει προσαρμοστεί για να υλοποιήσει τη συγκεκριμένη διεύθυνση εύρους ζώνης.

Η τοποθέτηση των εργασιών στα άκρα του δικτύου γίνεται τυχαία και συνδυάζεται με διάφορες εφαρμογές, προσομοιώνοντας τα αποτελέσματα του κατακερματισμού όταν πολλές διαφορετικές εφαρμογές μοιράζονται το δίκτυο. Το δίκτυο προ-διαμορφώνεται για κάθε προσομοίωση, και στη συνέχεια επιδεικνύουμε στην πειραματική μας εγκατάσταση τον έλεγχο που απαιτείται για την αναδιαμόρφωση.

Τα μοντέλα ισχύος και επιδόσεων σε επίπεδο συστήματος περιλαμβάνουν 16×16 φωτονικούς μεταγωγείς PINE και ηλεκτρικούς δρομολογητές Mellanox InfiniBand 100 Gb/s με 36 θύρες, τα οποία συνδέονται με ενεργά οπτικά καλώδια και έχουν διαστάσεις που αντιστοιχούν στα ιχνηλατημένα δεδομένα της εφαρμογής.

Στο Σχήμα 4.3 συνοψίζονται οι βελτιώσεις στην απόδοση και την καθυστέρηση από τη διεύθυνση εύρους ζώνης. Όπως φαίνεται, κατά μέσο όρο, η καθοδήγηση εύρους ζώνης βελτιώνει τη διαρκή μέση απόδοση του δικτύου κατά 1,7 φορές και μειώνει τη μέση καθυστέρηση κατά 20%. Οι βελτιώσεις αυτές δείχνουν ότι η καθοδήγηση εύρους ζώνης αποκαθιστά την απόδοση που χάνεται στην βασική τοπολογία λόγω μη ιδανικής εξισορρόπησης φορτίου ανά πακέτο.

Επιπλέον, η καθοδήγηση εύρους ζώνης βελτιώνει την απόδοση με οικονομικό τρόπο, χωρίς να απαιτείται υπερ-σχεδίαση για τις χειρότερες περιπτώσεις, καθώς οποιαδήποτε βασική γραμμή μπορεί να προσφέρει πλήρη ρυθμό μετάδοσης, αυξάνοντας το εύρος ζώνης των συνδέσεων και των μεταγωγέων. Τέλος, το εύρος ζώνης μεγιστοποιεί τα οφέλη της διεύθυνσης εύρους ζώνης.



Σχήμα 4.3 Μέση βελτίωση της απόδοσης του συστήματος του κατευθυνόμενου δέντρου σε σύγκριση με ένα vanilla fat tree χωρίς μείωση του εύρους ζώνης του ανώτερου στρώματος[5]

Η υψηλότερη απόδοση του δικτύου σημαίνει ότι το σύστημα μπορεί να διαχειριστεί περισσότερες συναλλαγές ανά δευτερόλεπτο, διατηρώντας σχεδόν την ίδια στατική κατανάλωση ισχύος (χρησιμοποιώντας τα ίδια στοιχεία δικτύου). Στην αρχιτεκτονική PINE, η αναδιαμόρφωση του οπτικού switch SiP συμβαίνει κάθε φορά που ξεκινά ή τερματίζεται μια εφαρμογή στο σύστημα.

Συνήθως, οι δείκτες αναφοράς HPC δημιουργούν μόνιμη κυκλοφορία με μοτίβα που αλλάζουν αργά. Για παράδειγμα, στο Εθνικό Κέντρο Ερευνών Ενέργειας (NERSC), οι εφαρμογές πολλαπλών κόμβων ξεκινούν κατά μέσο όρο κάθε 17 δευτερόλεπτα, πολύ πιο αργά από τις δυνατότητες του οπτικού μας SiP, το οποίο αναδιαμορφώνεται σε 20 μs. Μια βασική διαπίστωση είναι ότι οι αποδοτικοί οπτικοί μεταγωγείς SiP επιβάλλουν αμελητέα δυναμική ισχύ και καθυστέρηση μόλις διαμορφωθούν.

Η κλιμάκωση του εύρους ζώνης που επιτρέπει το PINE στο ανώτερο επίπεδο δεν προκαλεί ποινές απόδοσης, αλλά αντίθετα, αυξάνει άμεσα τις συναλλαγές ανά δευτερόλεπτο για τον ίδιο όγκο ισχύος. Τα αποτελέσματα από τη φάση 1 δείχνουν συνολικό μέσο παράγοντα βελτίωσης 2,2 φορές στις συναλλαγές ανά τζάουλ (55% μείωση της κατανάλωσης ενέργειας ανά σταθερού μεγέθους συναλλαγές) σε ένα ποικίλο σύνολο εφαρμογών HPC και κέντρων δεδομένων.

Περαιτέρω οφέλη προκύπτουν για εφαρμογές που έχουν περιορισμένο εύρος ζώνης επικοινωνίας, καθώς η καθοδήγηση του εύρους ζώνης επιταχύνει άμεσα την εκτέλεση της εφαρμογής κατά μέσο όρο κατά 1,7 φορές, γεγονός που μειώνει την αδρανή ισχύ υπολογιστών κατά τον ίδιο παράγοντα.

Με βάση τα επιτυχημένα αποτελέσματα της φάσης 1, οι προσπάθειες της φάσης 2 του προγράμματος PINE θα επικεντρωθούν σε διάφορους κρίσιμους τομείς. Αυτά περιλαμβάνουν:

1. Προηγμένος αναδιαμορφώσιμος σχεδιασμός δικτύου: Η φάση 2 θα αξιοποιήσει τα οφέλη ενός πιο προχωρημένου σχεδιασμού δικτύου, επιτρέποντας μεγαλύτερη ευελιξία και προσαρμοστικότητα στις μεταβαλλόμενες απαιτήσεις των εφαρμογών.
2. Ανασυγκρότηση πολλαπλών φορτίων εργασίας: Θα εξεταστεί η δυνατότητα διαχείρισης και προσαρμογής σε διάφορα φορτία εργασίας ταυτόχρονα, βελτιστοποιώντας τη χρήση των πόρων του συστήματος.
3. Πιο λεπτομερής χρονική αναδιαμόρφωση: Οι διαδικασίες αναδιαμόρφωσης του δικτύου θα γίνουν πιο λεπτομερείς, επιτρέποντας γρηγορότερες και πιο ακριβείς προσαρμογές στις απαιτήσεις του συστήματος.
4. Διαφορετική φυσική συνδεσιμότητα: Θα εξεταστούν και θα αξιολογηθούν διάφορες φυσικές συνδέσεις μεταξύ ηλεκτρονικών και οπτικών μεταγωγέων SiP στο δοκιμαστικό σύστημα PINE, προκειμένου να βελτιωθούν οι επιδόσεις και η αποδοτικότητα του συστήματος.
5. Διεύθυνση εύρους ζώνης: Η έννοια της διεύθυνσης εύρους ζώνης θα εμβαθύνει περαιτέρω στο διαχωρισμένο ράφι (disaggregated rack), επιδιώκοντας τη βελτιστοποίηση της κυκλοφορίας και της συνδεσιμότητας στο σύστημα.

Η φάση 2 στοχεύει να ενισχύσει περαιτέρω την αποδοτικότητα και τις επιδόσεις του PINE, προσδιορίζοντας και υλοποιώντας καινοτόμες προσεγγίσεις για τη διαχείριση του εύρους ζώνης και της συνδεσιμότητας.

Φάση 2 του PINE: Αρχιτεκτονική για βαθύ διαχωρισμό(Deep Disaggregation)

Ο αρχιτεκτονικός σχεδιασμός της φάσης 2 του προγράμματος PINE (Σχήμα 4.4) εισάγει μια καινοτόμο προσέγγιση για την αναδιάρθρωση των παραδοσιακών υπολογιστικών κέντρων δεδομένων και HPC (High Performance Computing). Οι κύριες αρχές του σχεδιασμού περιλαμβάνουν:

1. Διαχωρισμός και Αναδιοργάνωση

Οι υπολογιστικοί κόμβοι και εξυπηρετητές διαχωρίζονται και αναδιοργανώνονται γύρω από ένα αναδιαμορφώσιμο ιστό δικτύου. Αυτός ο σχεδιασμός έχει στόχο την αποτελεσματική διαχείριση της πίεσης που ασκούν οι εφαρμογές έντασης επικοινωνίας σε πραγματικό χρόνο.

2. Ενσωματωμένη Φωτονική Μεταγωγή

Το δίκτυο διασύνδεσης περιλαμβάνει ενσωματωμένες φωτονικές μεταγωγές, που επιτρέπουν τη διαχείριση του εύρους ζώνης ανάλογα με τις απαιτήσεις των διαφόρων πόρων. Η κατεύθυνση του εύρους ζώνης γίνεται σύμφωνα με τις ανάγκες, επιτρέποντας την αποδοτικότερη επικοινωνία.

3. Υψηλή Πυκνότητα και Ενεργειακή Αποδοτικότητα

Ο βαθύς διαχωρισμός επιτυγχάνεται μέσω της συναρμολόγησης εξαιρετικά υψηλής πυκνότητας ενεργειακά αποδοτικών φωτονικών συνδέσεων με GPU, CPU και στοιχεία μνήμης. Αυτές οι συνδέσεις τοποθετούνται σε μια πλατφόρμα MCM (Multi-Chip Module) interposer, που περιβάλλει ένα ενοποιημένο και αναδιαμορφώσιμο δίκτυο φωτονικών συνδέσεων.

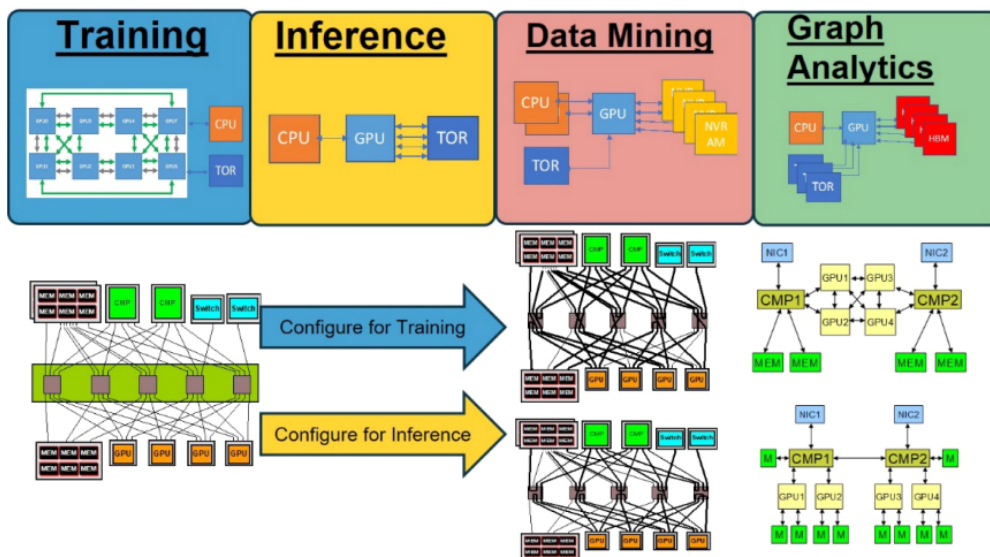
4. Ευέλικτη Διασυνδεσιμότητα

Το PINE επιτρέπει την ευέλικτη κατανομή πόρων σε φόρτους εργασίας με εξαιρετική χρονική και ποσοτική ακρίβεια. Η κατανομή των πόρων γίνεται σύμφωνα με τις απαιτήσεις των εφαρμογών, εξασφαλίζοντας ότι παρέχεται μόνο η απαραίτητη υπολογιστική ισχύς, χωρητικότητα μνήμης και εύρος ζώνης διασύνδεσης για την απαιτούμενη χρονική περίοδο.

5. Μείωση Σπατάλης Ενέργειας και Βελτίωση Απόδοσης

Αυτή η αποδοτική χρήση των πόρων οδηγεί σε σημαντική μείωση της σπατάλης ενέργειας που παρατηρείται στα σημερινά κέντρα δεδομένων. Η βελτιωμένη αποδοτικότητα αυξάνει την απόδοση της επένδυσης και επιταχύνει το χρόνο ολοκλήρωσης των εφαρμογών HPC, καθώς η αποτελεσματική επικοινωνία μεταξύ των πόρων βελτιώνει τη συνολική απόδοση του συστήματος.

Η φάση 2 του PINE στοχεύει στη δημιουργία ενός πιο ευέλικτου, αποδοτικού και ενεργειακά βιώσιμου δικτύου, που μπορεί να προσαρμοστεί στις απαιτήσεις των σύγχρονων εφαρμογών HPC. Η καινοτόμος αυτή προσέγγιση έχει τη δυνατότητα να προσφέρει σημαντικές βελτιώσεις στην απόδοση και την αποδοτικότητα των υποδομών κέντρων δεδομένων.



Σχήμα 4.4 Σχηματική απεικόνιση της βαθιάς διάσπασης και των προσαρμοσμένων διαμορφώσεων κόμβων για την υποστήριξη διαφορετικών φόρτων εργασίας[5]

Οι φόρτοι εργασίας στα κέντρα δεδομένων και τα συστήματα HPC (High Performance Computing) παρουσιάζουν μεγάλη ποικιλία στις απαιτήσεις τους σε πόρους. Αυτή η ποικιλία επιβάλλει νέες αρχιτεκτονικές σχεδίασης που μπορούν να ανταποκριθούν στις ιδιαίτερες ανάγκες κάθε φόρτου εργασίας. Ακολουθούν ορισμένα κρίσιμα σημεία σχετικά με τις προκλήσεις και την προσέγγιση της αρχιτεκτονικής PINE:

1. Ποικιλία Φορτίων Εργασίας

- Εκπαίδευση Αλγορίθμων: Οι διαδικασίες εκπαίδευσης αλγορίθμων για βαθιά μάθηση απαιτούν σημαντική υπολογιστική ισχύ και επικοινωνία, οδηγώντας σε άκαμπτα πρότυπα επικοινωνίας.
- Βάσεις Δεδομένων στη Μνήμη: Αυτές οι βάσεις καταπονούν τη μνήμη και το εύρος ζώνης αποθήκευσης, απαιτώντας γρήγορη και αποδοτική πρόσβαση στα δεδομένα.
- Αναλύσεις Έντασης Δεδομένων: Αυτές οι εργασίες απαιτούν υψηλή χωρητικότητα και εύρος ζώνης μνήμης για να μπορούν να διαχειριστούν τα μεγάλα σύνολα δεδομένων αποτελεσματικά.

2. Σύγχρονες Αρχιτεκτονικές

Πολλές από τις υπάρχουσες αρχιτεκτονικές διαχωρίζουν έντονα τους πόρους, όπως την αποθήκευση, τη διασύνδεση και τη μνήμη, σε μεμονωμένους διακομιστές. Αυτή η απομόνωση των πόρων εμποδίζει την ευελιξία και την ανταλλαγή πόρων μεταξύ τους, οδηγώντας σε υποχρεωτική προμήθεια υπερβολικών στοιχείων του συστήματος.

3. Αποτελεσματικότητα και Ασυμμετρία

Οι άκαμπτες αρχιτεκτονικές δεν ανταποκρίνονται στις μεταβαλλόμενες απαιτήσεις των φόρτων εργασίας, με αποτέλεσμα τη σπατάλη ενέργειας και πόρων. Οι υπερβολικές προμήθειες οδηγούν σε αναποτελεσματική χρήση και υποχρησιμοποίηση πόρων, αντί να βελτιστοποιούνται σύμφωνα με τις ανάγκες των εφαρμογών.

4. Προτεινόμενη Λύση: Αρχιτεκτονική PINE

Η αρχιτεκτονική PINE προτείνει τη χρήση πολυάριθμων, στρατηγικά διατεταγμένων οπτικών κυκλωμάτων χαμηλού έως μεσαίου radix, που κατευθύνουν το εύρος ζώνης σύμφωνα με τη ζήτηση. Αυτή η προσέγγιση επιτρέπει την παροχή βελτιστοποιημένης συνδεσιμότητας σε εφαρμογές, μειώνοντας έτσι τη σπατάλη ενέργειας από ανενεργούς πόρους.

5. Απαιτήσεις και Πλεονεκτήματα

Η αρχιτεκτονική PINE απαιτεί μόνο μεταγωγείς χαμηλής έως μεσαίας ακτίνας, οι οποίοι προσφέρουν χαμηλές απώλειες και κατανάλωση ισχύος. Αυτό επιτρέπει την αποτελεσματική διαχείριση του εύρους ζώνης και τη βελτίωση της απόδοσης σε όλα τα επίπεδα του συστήματος.

Βαθύς διαχωρισμός σε επίπεδο κόμβου (Node Level Deep Disaggregation)

Τα τρέχοντα σχέδια τσιπ για διακομιστές είναι σταθερά συνδεδεμένα με συγκεκριμένες πηγές. Ως αποτέλεσμα, σχεδόν δεν παρέχουν καμία δυνατότητα αναδιάρθρωσης του εύρους ζώνης εισόδου/εξόδου και μνήμης για την ικανοποίηση των απαιτήσεων των εφαρμογών. Για παράδειγμα, οι εφαρμογές μηχανικής μάθησης χρειάζονται τουλάχιστον 3:1 μετατόπιση του εύρους ζώνης εισόδου/εξόδου και μνήμης από τη σύνδεση μεταξύ GPU (για εκπαίδευση) σε εύρος ζώνης εκτός τσιπ (για την εξαγωγή συμπερασμάτων). Αν δεν γίνει αυτό, θα προκύψει σημαντική ποινή στην ενεργειακή αποδοτικότητα, είτε μέσω μειωμένης απόδοσης της εφαρμογής είτε μέσω χαμηλότερης παροχής εύρους ζώνης.

Η προσέγγιση της βαθιάς διάσπασης του PINE θα επιτύχει αυτήν τη μετατόπιση 3:1 στην ισορροπία εύρους ζώνης/σύνδεσης και θα αναπτύξει νέους αλγορίθμους για να καθοδηγήσει την αναδιάρθρωση του φωτονικού δικτύου μεταγωγής MCM, με σκοπό τη διαμόρφωση προσαρμοσμένων κόμβων από απομονωμένους πόρους για να ικανοποιήσουν διαφορετικές απαιτήσεις των εφαρμογών. Οι εφαρμογές βιοπληροφορικής και γραφημάτων απαιτούν ακόμα μεγαλύτερες μετατοπίσεις, 10:1 ή και περισσότερες, στην παροχή εύρους ζώνης για την εκπλήρωση των απαιτήσεων τους.

Στη δεύτερη φάση, το PINE θα παρέχει δυναμικό εύρος 10:1 ή μεγαλύτερο για την αναδιάρθρωση και επανασύνδεση των κόμβων, το οποίο με τη σειρά του θα προσφέρει ένα πλεονέκτημα απόδοσης άνω του 5× ανά Watt για εφαρμογές με ποικιλία απαιτήσεων πόρων. Αυτά τα κέρδη προσδιορίζονται με τη χρήση της μετρικής TEPS ανά Watt.

Ο φωτονικός κόμβος MCM PINE επιτρέπει με μοναδικό τρόπο τη δημιουργία ενός ευέλικτου «φωτονικού ιστού» που μπορεί να ανασυνδέει στρατηγικά αποσπασματικά στοιχεία για να σχηματίσει «κόμβους» κατά τη διάρκεια της λειτουργίας, ανταγωνιζόμενος έτσι διαφορετικές απαιτήσεις φόρτου εργασίας. Πραγματικά, καταργεί τα όρια μεταξύ του κόμβου και του ραφιού. Στη δεύτερη φάση, θα επικεντρωθούμε στην ενίσχυση της ευελιξίας του εύρους ζώνης. Το Σχήμα 4.4 απεικονίζει σχηματικά πώς οι τέσσερις τυπικοί φόρτοι εργασίας διαθέτουν διαφορετικές οργανώσεις κόμβων για να υποστηρίξουν τις απαιτήσεις τους. Ο βαθύς διαχωρισμός επιτρέπει τη διαμόρφωση προσαρμοσμένων διατάξεων κόμβων σε διαδικασία εκκίνησης, υποστηρίζοντας αυτές τις ποικιλόμορφες απαιτήσεις φόρτου εργασίας με βελτιστοποιημένη ενεργειακή συνδεσιμότητα.

Η δεύτερη φάση θα επικεντρωθεί σε μια αλγοριθμική προσέγγιση για τη ρύθμιση και βελτιστοποίηση της ενεργειακής απόδοσης των φορτίων εργασίας. Γίνεται ανάπτυξη στην ευέλικτη συνδεσιμότητα της φωτονικής μεταγωγής εντός της αρχιτεκτονικής του κόμβου, καθώς και περιγραφή του επιπέδου ελέγχου που απαιτείται για να επιτραπεί η διαμόρφωση κατά τη διάρκεια εκτέλεσης με κλιμακούμενο τρόπο εντός ενός ραφιού. Για παράδειγμα, οι διάφορες φάσεις των εφαρμογών εκμάθησης μπορούν να συναρμολογούνται δυναμικά ώστε να παρέχουν το απαιτούμενο εύρος ζώνης συνδεσιμότητας μεταξύ GPU και μνήμης.

Μια κοινή προσέγγιση στην καταναμημένη εκπαίδευση είναι ο παραλληλισμός δεδομένων, όπου τα δεδομένα εκπαίδευσης κατανομούνται σε πολλαπλούς επεξεργαστές (όπως GPU, TPU, CPU). Στην παράλληλη εκπαίδευση δεδομένων, οι επεξεργαστές πρέπει να επικοινωνούν τις παραμέτρους του μοντέλου τους μετά από κάθε επανάληψη εκπαίδευσης. Αυτό μπορεί να γίνει με διάφορες μεθόδους, όπως μέσω διακομιστών παραμέτρων, ring-allreduce, tree-reduce και ιεραρχικού all-reduce. Ωστόσο, υπάρχει ραγδαία αύξηση των

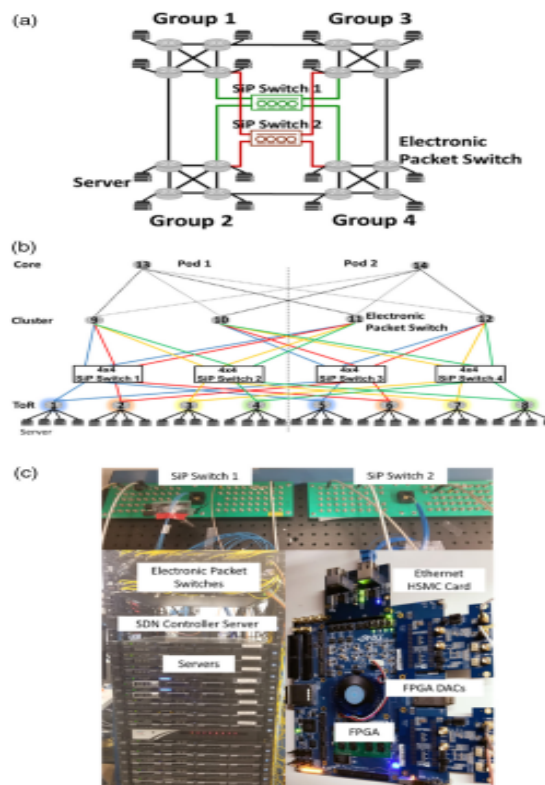
μοντέλων και της εκπαίδευσης με παραλληλισμό αγωγών, λόγω της ταχείας αύξησης των υπολογιστικών και μνημονικών απαιτήσεων των νευρωνικών δικτύων.

Για την εκπαίδευση μοντέλων, ο παραλληλισμός μοντέλων (μαζί με υβριδικές προσεγγίσεις παραλληλισμού δεδομένων-μοντέλου) διασπά το μοντέλο (και τα δεδομένα) σε διαφορετικούς επεξεργαστές. Τέλος, ο παραλληλισμός μοντέλου είναι ένα ενεργό πεδίο έρευνας, με διάφορες τεχνικές καταμερισμού μοντέλων.

Εφαρμογή του πειραματικού συστήματος PINE

Για να γίνει εξέταση της σκοπιμότητας της έννοιας διεύθυνσης εύρους ζώνης PINE μέσω διακοπών SiP, έχει δημιουργηθεί ένα δοκιμαστικό περιβάλλον HPC/κέντρου δεδομένων. Αυτό το περιβάλλον ενσωματώνει τους switch SiP με ένα παραδοσιακό ηλεκτρονικό σύστημα μεταγωγής πακέτων, το οποίο αποτελείται από εμπορικούς διακομιστές και μεταγωγείς πακέτων, καθώς και προγραμματιζόμενες στο πεδίο πύλες (FPGAs) και δικτύωση καθορισμένη από λογισμικό (SDN). Στο περιβάλλον αυτό, μπορούμε να εκτελούμε οποιαδήποτε λειτουργία ενός συμβατικού υπολογιστικού συστήματος, περιλαμβάνοντας μαζική μεταφορά δεδομένων, εικονικές μηχανές (VM) και εφαρμογές αναφοράς HPC.

Με την ενσωμάτωση των μεταγωγέων SiP στο δίκτυο και την εφαρμογή της διεύθυνσης εύρους ζώνης για τη βελτιστοποίηση της δικτυακής τοπολογίας, η φάση 1 κατέδειξε σημαντικές βελτιώσεις στην απόδοση, μέσω της μείωσης του συνολικού χρόνου εκτέλεσης της εφαρμογής σε σύγκριση με παραδοσιακές δικτυακές τοπολογίες, γεγονός που μεταφράζεται σε αντίστοιχη βελτίωση της ενεργειακής κατανάλωσης. Το δοκιμαστικό περιβάλλον HPC/κέντρου δεδομένων PINE με 32 κόμβους (Σχ. 4.5) περιλαμβάνει τέσσερις μεταγωγείς SiP 4×4 και μπορεί να διαμορφωθεί σε τοπολογία dragonfly ή σε τοπολογία fat tree.



Σχήμα 4.5 Φυσικά δίκτυα δοκιμών: (a) dragonfly, (b) fat tree και (c) implemented hardware[5]

Το δοκιμαστικό σύστημα της φάσης 1 του PINE αξιολόγησε το εύρος ζώνης της εφαρμογής αναφοράς GTC HPC και παρουσίασε σημαντική βελτίωση στις επιδόσεις, επιταχύνοντας τον χρόνο εκτέλεσης κατά περισσότερο από 62% σε σχέση με το τυπικό fat tree. Στη φάση 2, το testbed θα επεκταθεί σημαντικά σε 64 κόμβους και θα προσαρμοστεί ώστε να αντικατοπτρίζει πιο ακριβώς τα πραγματικά αναπτυγμένα υπολογιστικά συστήματα. Μόλις μια εφαρμογή τοποθετηθεί στους φυσικούς κόμβους, αναμένουμε ότι η βέλτιστη συνδεσιμότητα θα αλλάξει με βάση το αναμενόμενο μοτίβο κίνησης.

Ως εκ τούτου, στη φάση 2, προτείνουμε την αναδιάρθρωση της τοπολογίας εντός δεκάδων δευτερολέπτων, που είναι ο μέσος χρόνος άφιξης για μια νέα εφαρμογή. Κατά τη διάρκεια κάθε περιόδου αναδιαμόρφωσης, ο ελεγκτής δικτύου θα εξάγει ένα αναμενόμενο μοτίβο κίνησης με βάση τις τρέχουσες τοποθετήσεις εργασιών στο δίκτυο και θα υλοποιεί τη λογική συνδεσιμότητα, έτσι ώστε η συνολική τοπολογία να βελτιστοποιείται για το αναμενόμενο μοτίβο κίνησης. Η εισαγωγή του PINE με οπτικά διασυνδεδεμένα MCM και φωτονική μεταγωγή θα πραγματοποιηθεί για την επικύρωση της διαδικασίας.

Το κλιμακωτό σύστημα θα επιτρέψει πιο ρεαλιστικά πειράματα κίνησης, αξιολογώντας την επίδραση της συμφόρησης της σύνδεσης, της καθυστέρησης και της απόδοσης σε ένα ευρύ φάσμα εφαρμογών HPC και μηχανικής μάθησης, με τη χρήση ενσωματωμένων GPU στους διακομιστές, καθώς και της κίνησης στο κέντρο δεδομένων.

4.3 Οπτικά διασυνδεδεμένα Multi-Chip Modules (MCM)

MCM Interposer

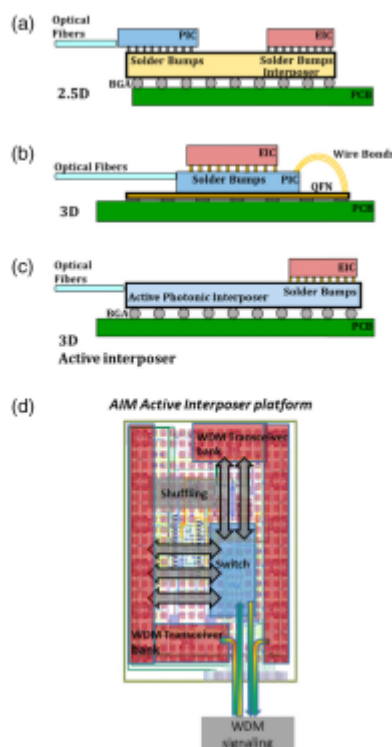
Στη φάση 1, αναπτύχθηκαν πομποδέκτες MCM για την αποτελεσματική ενσωμάτωση της φωτονικής με τα ηλεκτρονικά συστήματα οδήγησης. Δημιουργήθηκαν διάφορα πρωτότυπα MCM σε 2,5D και 3D πλατφόρμες παρεμβολής [Σχήματα 4.6(α) και 4.6(β)]. Η ενεργή πλατφόρμα παρεμβολής, που αναπτύχθηκε σε συνεργασία με το SUNY Poly, συνδυάζει μοναδικά το φωτονικό ολοκληρωμένο κύκλωμα (PIC) με το interposer σε ένα ενιαίο ολοκληρωμένο υπόστρωμα [Σχ. 4.6(γ)]. Αυτή η προσέγγιση μπορεί να προσφέρει τις καλύτερες ηλεκτρονικές και φωτονικές πυκνότητες και αναμένεται να αποτελέσει τον κύριο δρόμο για τη φάση 2, ολοκληρώνοντας τη σύνδεση φωτονικής και τον switch με τα προσαρμοσμένα ηλεκτρονικά ολοκληρωμένα κυκλώματα (EIC).

Η μετάβαση από το 2,5D στο 3D ενεργό interposer θα είναι μια κοινή προσπάθεια μεταξύ των Columbia/SUNY Poly και NVIDIA. Ο ενεργός interposer συνδυάζει τις λειτουργίες του PIC και του interposer σε ένα ενιαίο στοιχείο, επιτρέποντας την κατασκευή και άμεση ενσωμάτωση φωτονικών στοιχείων μέσω διανοιγμάτων στο πυρίτιο (TSVs) και πρόσθετα στρώματα ανακατανομής μετάλλων για να επιτρέψει τη συνδεσιμότητα τόσο στην μπροστινή όσο και στην πίσω πλευρά του ενεργού interposer.

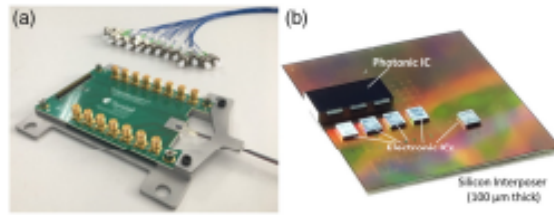
Εκτός από τα στρώματα παρεμβολής, ο ενεργός interposer θα περιλαμβάνει όλα τα χαρακτηριστικά που βρίσκονται στο PIC, διευκολύνοντας την κατασκευή και τη δρομολόγηση ενεργών και παθητικών φωτονικών διατάξεων. Η επάνω πλευρά του ενεργού παρεμβύσματος θα χρησιμοποιηθεί για την παροχή συνδεσιμότητας με τα κυκλώματα οδήγησης συμπληρωματικού οξειδίου μετάλλου-ημιαγωγού (CMOS) EIC. Τα EIC θα

τοποθετηθούν στην κορυφή του ενεργού interposer χρησιμοποιώντας χάλκινους πυλώνες. Οι πυλώνες αυτοί θα αναπτυχθούν σε κλίμακα πλακιδίου τόσο στο ενεργό interposer όσο και στα EIC.

Η επάνω πλευρά του ενεργού interposer θα προσφέρει επίσης μια πλατφόρμα για ενσωμάτωση σε υπολογιστικούς κόμβους, όπως CPU, GPU, μνήμη ή chip VLSI. Η πίσω πλευρά του ενεργού interposer θα προορίζεται για σφραγίδες (BGA) με την πλακέτα τυπωμένου κυκλώματος (PCB). Στο σχήμα 4.6(δ) απεικονίζεται ένα παράδειγμα δικτύου 8×8 σε τσιπ που υλοποιείται στην πλατφόρμα ενεργού interposer. Το πλήρως πακεταρισμένο πρωτότυπο του πομποδέκτη MCM, τοποθετημένο στην πλακέτα PCB με ανεμιστήρα, απεικονίζεται στο Σχήμα 4.7(α), ενώ ο πομποδέκτης MCM που χρησιμοποιεί παθητικό παρεμβαλλόμενο πυρίτιο για ολοκλήρωση 2,5D φαίνεται στο Σχήμα 4.7(β).



Σχήμα 4.6 (α)-(γ) Προσεγγίσεις 2,5D και 3D MCM που διερευνήθηκαν στη φάση 1- δ) δίκτυο 8×8 στο τσιπ που υλοποιήθηκε στο ενεργό πλατφόρμα interposer.[5]



Σχήμα 4.7 (α) Πλήρως συσκευασμένος πρωτότυπος πομποδέκτης MCM (β) Πομποδέκτης MCM που χρησιμοποιεί παθητικό πυρίτιο interposer για ενσωμάτωση 2,5D.[5]

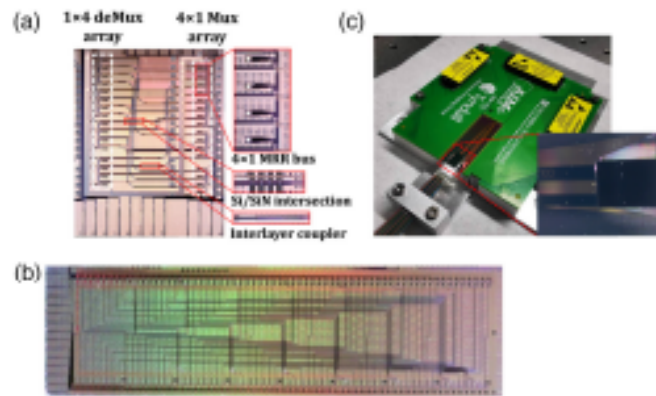
MCM Photonic Switch Fabric

Η ενσωμάτωση του δικτύου φωτονικών διακοπών στην πλατφόρμα MCM επιτρέπει την αναδιαμόρφωση του οπτικού τομέα για τη βαθιά διάσπαση του PINE. Στη φάση 1, αξιοποιήσαμε την πλατφόρμα ενεργού παρεμβολής AIM και αναπτύξαμε το πρώτο πρωτότυπο ενός οπτικού δικτύου σε τσιπ, το οποίο συνδυάζει μονολιθικά έναν χωρικό switch 8×8 με τέσσερις πομπούς WDM, τέσσερις δέκτες WDM, και EICs τοποθετημένα από πάνω (Σχήμα 4.6(δ)).

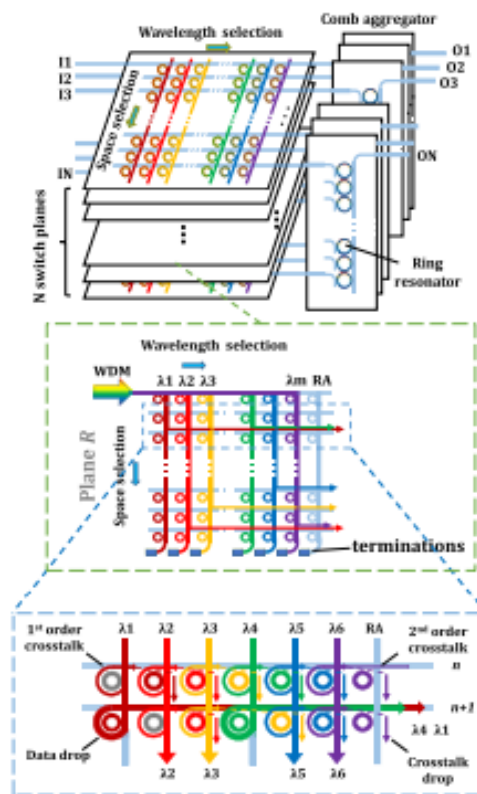
Στη φάση 1 κατασκευάστηκε ο πρώτος αυστηρά μη-μπλοκαρισμένος οπτικός switch σε τοπολογία switch-and-select (S&S), χρησιμοποιώντας αντηχεία microring [9,10], όπως φαίνεται στο Σχήμα 4.8(α). Η μοναδική πολυστρωματική τοπολογία S&S από Si/SiN ανέδειξε πρωτοποριακή καταστολή της διασταυρούμενης ομιλίας του switch, με λόγο εξάλειψης >50 dB και απώλεια στο τσιπ ως $<1,8$ dB. Έχει επίσης κατασκευαστεί ένα τριπλό στρώμα 8×8 μικροδιακόπτη(micro-switch) στην ίδια τοπολογία, με μικροσκοπική εικόνα που παρουσιάζεται στο Σχήμα 4.8(β). Το τσιπ του switch συσκευάστηκε χρησιμοποιώντας παρεμβαλλόμενο πυρίτιο, που μπορεί εύκολα να ενσωματωθεί στην πλατφόρμα MCM.

Στη φάση 2, γίνεται επέκταση της τοπολογίας για να υποστηριχθεί $>100 \times 100$ συνδεσιμότητα με αντηχεία microring και θα αναπτυχθεί καθοδήγηση εύρους ζώνης τόσο στο χωρικό όσο και στο φασματικό τομέα για μεγαλύτερη ευελιξία [26]. Η αρχιτεκτονική του switch περιλαμβάνει δύο τμήματα: 1) N επίπεδα μεταγωγής ενός $N \times M$ λ πίνακα διασταυρώσεων για χωρική επιλογή μήκους κύματος και 2) N συσσωρευτές μήκους κύματος comb ελεύθερης φασματικής περιοχής (FSR), όπως απεικονίζεται στο Σχήμα 4.9.

Κάθε επίπεδο μεταγωγής $N \times M$ λ αποτελείται από έγχρωμες συστοιχίες μικροδακτυλίων σε N σειρές και $(M + 1)$ στήλες, όπου η επιλογή μήκους κύματος γίνεται από την πρώτη σειρά, ενώ η χωρική επιλογή ενεργοποιείται από τους δακτυλίους σε κάθε στήλη. Οι comb συσσωρευτές εφαρμόζουν στοιχεία μεγάλων δακτυλίων με προσαρμογή FSR. Αυτή η διάταξη θα επιτρέψει έναν κλιμακούμενο ιστό μεταγωγής που συνδυάζει τη μεταγωγή στον χωρικό τομέα με την επιλεκτικότητα μήκους κύματος, ώστε να καθορίζεται λεπτομερής συνδεσιμότητα για τη διάσπαση των κόμβων τόσο στη φυσική θύρα όσο και στο κανάλι μήκους κύματος.



Σχήμα 4.8 (α) Φωτογραφία μικροσκοπίου της κατασκευασμένης διάταξης με ένθετα του διευρυμένου χωρικού πολυπλέκτη 4 × 1 με βάση το MRR, των διασταυρώσεων Si/SiN και του διαστρωματικού ζεύκτη. (β) Φωτογραφία μικροσκοπίου του βιδωμένου switch η 8 × 8 τριπλών στρωμάτων, (γ) Πακεταρισμένος switch με παρεμβαλλόμενο υλικό πυριτίου.[5]



Σχήμα 4.9 Προτεινόμενος σχεδιασμός switch χώρου και μήκους κύματος με χρήση συστοιχιών επιλογών μήκους κύματος με βάση το microring και συσσωρευτών χτενίσματος. Δείχνουν την αρχή λειτουργίας του πίνακα διασταυρούμενων σημείων $N \times M\lambda$ και πώς μπλοκάρει πλήρως τη διασταυρούμενη ομιλία πρώτης τάξης εντός της ζώνης.[5]

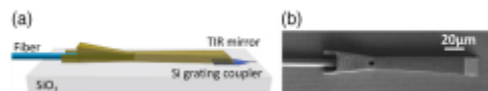
Ηλεκτρονικά στοιχεία εξαιρετικά χαμηλής ενέργειας

Οι πιο ενεργειακά αποδοτικοί ηλεκτρικοί δίαυλοι επικοινωνίας είναι αυτοί που μπορούν να υλοποιηθούν με τις απλούστερες αρχιτεκτονικές και κυκλώματα. Γι' αυτό το λόγο, οι αρχιτεκτονικές που προάγουν τη μετάδοση ομαδοποιημένων δεδομένων και ρολογιού χρησιμοποιούνται σχεδόν καθολικά σε ηλεκτρικές διασυνδέσεις μικρής απόστασης που απαιτούν αποδοτικότητα στην κατανάλωση ενέργειας. Στην αρχιτεκτονική του συστήματός μας, θα εισάγουμε ένα κοινό ρολόι για κάθε ομάδα οκτώ οπτικών καναλιών δεδομένων και θα αξιοποιήσουμε αυτό το ρολόι για τη δειγματοληψία των δεδομένων στον δέκτη, αποφεύγοντας έτσι την πολύπλοκη διανομή ρολογιού και τα κυκλώματα ανάκτησής του. Έχουμε μελετήσει σε βάθος παρόμοιες τεχνικές προώθησης ρολογιού για ηλεκτρικούς συνδέσμους μικρών αποστάσεων.

Αυτή η τεχνική προώθησης ρολογιού διευκολύνει την ανάπτυξη πολύ απλού κυκλώματος μετάδοσης, στο οποίο τα δεδομένα είναι σειριακά, χρονίζονται με ένα κοινό ρολόι και μεταδίδονται στον οδηγό του διαμορφωτή.

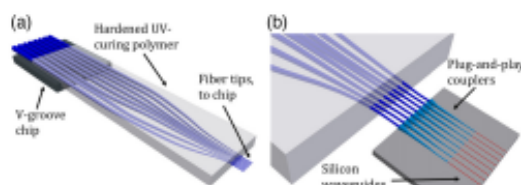
Ανθεκτικό οπτικό ΙΟ υψηλής πυκνότητας

Η τρισδιάστατη δομή που φαίνεται στο Σχήμα 4.10, περιλαμβάνει ένα πολυμερές χωνί που κατευθύνει την εισερχόμενη ίνα (αραιωμένη στο ένα άκρο) απευθείας προς την όψη ενός πολυμερούς κυματοδηγού. Η ίνα δρομολογείται στη χοάνη ανεξάρτητα από την ακριβή θέση της σε σχέση με το κέντρο της (διάμετρος 20 μm), παρέχοντας έτσι εξαιρετικά υψηλή ανοχή ευθυγράμμισης. Όταν η ίνα εισάγεται πλήρως, συγκρατείται σφιχτά στο μικρότερο άκρο του τμήματος δρομολόγησης της χοάνης. Σε αυτό το στάδιο, η ίνα και ο πολυμερής κυματοδηγός είναι τέλεια ευθυγραμμισμένοι, οπτικά συζευγμένοι και μηχανικά συγκρατημένοι στη θέση τους. Οι οπτικοί τρόποι λειτουργίας τους έχουν σχεδιαστεί ώστε να ταιριάζουν χωρικά, εξασφαλίζοντας υψηλή απόδοση σύζευξης. Η κατασκευή του συζεύκτη πραγματοποιείται μέσω ενός τρισδιάστατου εργαλείου νανοεκτύπωσης δύο φωτονίων, που χρησιμοποιείται για τον πολυμερισμό ενός φωτοανθεκτικού υλικού βάσει εποξειδικής ρητίνης πάνω από τα κατασκευασμένα τσιπ κυματοδηγού. Οι τυπικές οπτικές ίνες λεπταίνουν στο ένα άκρο τους στα 10 μm (μια διαδικασία διαθέσιμη και στο εμπόριο) και εισάγονται απευθείας στη χοάνη (funnel).



Σχήμα 4.10 (a) Σχηματική απεικόνιση του ανθεκτικού link plug-and-play. (b) εικόνα ηλεκτρονικής μικροσκοπίας σάρωσης των κατασκευασμένων συσκευιών[5]

Η εργασία με αυτές τις λεπτές ίνες επιτρέπει την υψηλή πυκνότητα συσκευασίας συζευκτών, χρησιμοποιώντας συστοιχίες χαμηλού αποτυπώματος (με βήμα έως 30 μm), όπως φαίνεται στο Σχήμα 4.11. Τα αποτελέσματά μας αποδεικνύουν ότι ο τρισδιάστατος συζευκτήρας διοχέτευσης επιφέρει ποινή μόλις 0,05 dB, υποδεικνύοντας ότι ο συνδετήρας plug-and-play έχει ελάχιστη επίδραση στην αποδοτικότητα σύζευξης της συσκευής, ενώ παράλληλα επιτρέπει τη συμβατότητα με τα πρότυπα αυτοματοποιημένων εργαλείων ευθυγράμμισης.



Σχήμα 4.11 (a) Σχηματικό διάγραμμα της διάταξης λεπτών ινών για οκτώ ίνες με βήμα 30 μm . (b) Σύζευξη συστοιχίας ινών σε συστοιχία οκτώ ζεύκτες plug-and-play γραμμένοι απευθείας στο τσιπ φωτονικής.[5]

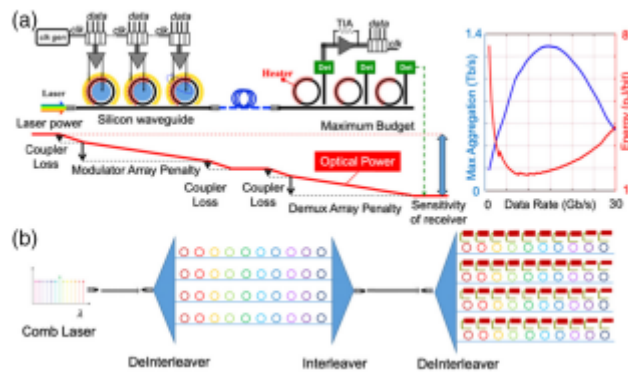
4.4 Φωτονικές συνδέσεις πυριτίου DWDM

Βελτιστοποιημένη Σχεδίαση Ενέργειας

Κατά την πρώτη φάση, αναπτύχθηκαν καινοτόμα εργαλεία προσομοίωσης και μοντελοποίησης των συνδέσεων PINE στο περιβάλλον PhoenixSim. Το PhoenixSim είναι σχεδιασμένο ώστε να βελτιστοποιεί την απόδοση του συστήματος, ελαχιστοποιώντας ταυτόχρονα την κατανάλωση ενέργειας. Μπορεί να χρησιμοποιηθεί για τον σχεδιασμό φωτονικών διατάξεων σε διάφορες παραμέτρους, όπως οι διασταυρώσεις με δακτυλιοειδή στοιχεία, οι ρυθμοί μετάδοσης bit καναλιών και οι μορφές διαμόρφωσης.

Πραγματοποιήθηκε λεπτομερή μοντελοποίηση πολλαπλών επιπέδων και ανάλυση ενεργειακής απόδοσης για συνδέσεις DWDM SiP, χρησιμοποιώντας διάφορα εξαρτήματα, όπως πηγές πολλαπλών μηκών κύματος, διαμορφωτές, φίλτρα, φωτοανιχνευτές και ηλεκτρικά ολοκληρωμένα στοιχεία (Εικ. 4.12(a)).

Ένα σημαντικό επίτευγμα της πλατφόρμας σχεδιασμού PhoenixSim στην πρώτη φάση ήταν η πλήρης ενσωμάτωση με το Synopsys Photonic Solutions, ένα εμπορικό λογισμικό για τον σχεδιασμό και την κατασκευή φωτονικών συστημάτων. Αυτή η ενσωμάτωση επιτρέπει έναν ευέλικτο σχεδιασμό, λαμβάνοντας υπόψη πολυφυσικά φαινόμενα στα φωτονικά εξαρτήματα, όπως η γεωμετρία της δομής, τα επίπεδα ντόπινγκ και η συχνότητα απόκρισης των διαμορφωτών.



Σχήμα 4.12 (a) Μοντέλα φωτονικής ζεύξης διαθέσιμα στο PhoenixSim εργαλείο λογισμικού. (b) Αρχιτεκτονική σύνδεσης DWDM εξαιρετικά χαμηλής κατανάλωσης με a πηγή λέιζερ comb. Οι γραμμές comb DWDM αποσυμπλέκονται σε τέσσερις ομάδες, και κάθε ομάδα έχει διαδοχικούς διαμορφωτές που βασίζονται σε μικροδακτυλίους.[5]

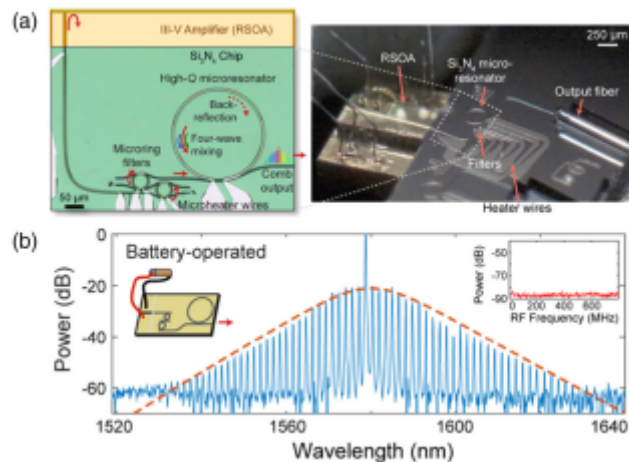
Η αρχιτεκτονική σύνδεσης για τη φάση 2, που παρουσιάζεται στο Σχ. 4.12(b), βασίζεται σε ένα comb DWDM που διαδίδεται μέσα στο τσιπ, όπου οι γραμμές comb χωρίζονται σε τέσσερις ομάδες, οι οποίες στη συνέχεια κατευθύνονται σε μια συστοιχία διαμορφωτών μικροδακτυλίου. Το κύκλωμα, το οποίο έχει σχεδιαστεί για εξαιρετικά χαμηλή κατανάλωση ενέργειας, αξιοποιεί τη σχέση ανάμεσα στην κατανάλωση ισχύος και την αναλογία μεταξύ του ρυθμού δεδομένων και της συχνότητας λειτουργίας των τρανζίστορ. Με τη μείωση του ρυθμού δεδομένων ανά κανάλι, επιτυγχάνεται σημαντική εξοικονόμηση ενέργειας, καθώς η κατανάλωση ισχύος αυξάνεται μη γραμμικά με τον ρυθμό δεδομένων. Αυτή η προσέγγιση επιτρέπει τη χρήση δεκτών με υψηλή ευαισθησία και προσφέρει τη δυνατότητα βελτιστοποίησης της ποιότητας του οπτικού σήματος. Ο προϋπολογισμός οπτικής σύνδεσης λαμβάνει υπόψη αναμενόμενες απώλειες σύζευξης, εισαγωγικές απώλειες εξαρτημάτων, παρεμβολές μεταξύ καναλιών WDM.

Η κατανάλωση ενέργειας του φωτονικού συνδέσμου εξαρτάται από την πηγή λέιζερ, τους διαμορφωτές, τους δέκτες και τα συναφή ηλεκτρονικά. Ο συνολικός προϋπολογισμός ισχύος βασίζεται στην αρχιτεκτονική που απεικονίζεται στο Σχ. 12(a). Υπάρχει μια συμβιβαστική σχέση μεταξύ του ρυθμού δεδομένων των επιμέρους διαμορφωτών και της συνολικής αρχιτεκτονικής του φωτονικού συνδέσμου για τη βέλτιστη ενεργειακή απόδοση. Για παράδειγμα, για να επιτευχθεί συνολικό εύρος ζώνης 640 Gb/s, ένας σύνδεσμος με διαμορφωτές κάτω από 10 Gb/s απαιτεί πιο σύνθετη αρχιτεκτονική, αυξάνοντας τις απώλειες διάδοσης. Αντίθετα, ένας σύνδεσμος με διαμορφωτές πάνω από 20 Gb/s απλοποιεί την αρχιτεκτονική, αλλά αυξάνει την ενέργεια που απαιτείται για διαμόρφωση και ανίχνευση.

Ενσωματωμένο λέιζερ comb DWDM

Οι πρόσφατες εξελίξεις στις comb συχνότητες που βασίζονται σε μικροσυντονιστές τσιπ ανοίγουν νέες δυνατότητες για την ελεγχόμενη παραγωγή πολλών στοιχείων συχνότητας. Αυτή η προσέγγιση είναι ιδανική για εφαρμογές WDM, καθώς η απόσταση μεταξύ των στοιχείων συχνότητας μπορεί εύκολα να καθοριστεί σε ένα συγκεκριμένο πλέγμα, σταθεροποιώντας τον μικροσυντονιστή.

Στην πρώτη φάση, επιτεύχθηκε 1) η επιτυχή ανάπτυξη μιας ενσωματωμένης πηγής comb, 2) η επίτευξη υψηλής απόδοσης αντλίας στη μετατροπή comb χρησιμοποιώντας την πλατφόρμα πυριτικού νιτριδίου (SiN), και 3) η πλήρης ενσωμάτωση σε μονάδα πομποδέκτη (Σχ. 4.13).



Σχήμα 4.13 (α) Εικόνα μικροσκοπίου μιας ενσωματωμένης πηγής comb (αριστερά) και φωτογραφία μιας πλήρως ενσωματωμένης πηγής comb (δεξιά) (β) Φάσμα comb Soliton που δημιουργείται με χρήση μπαταρίας AAA παρέχοντας 98 mW ηλεκτρικής ενέργειας.[5]

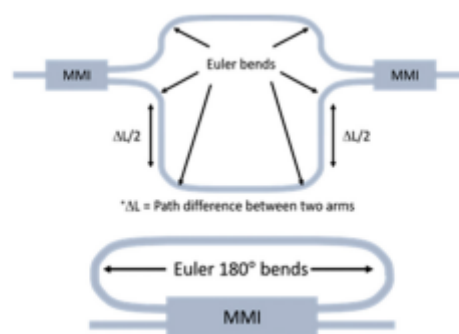
Ισχυρά φίλτρα DWDM

Στην αρχιτεκτονική αυτή, τα δέντρα φίλτρων που βασίζονται σε συμβολόμετρα Mach-Zehnder (MZI) θα συνδυαστούν με αντηχεία δακτυλίου με στενό εύρος Ελεύθερης Φασματικής Περιοχής (FSR), τα οποία θα φιλτράρουν περαιτέρω το σήμα για να απομονώσουν το επιθυμητό μήκος κύματος και εύρος ζώνης. Σήμερα, οι κυματοδηγοί πυριτίου με υψηλό περιορισμό και μικρές διαστάσεις, αν και αποτελεσματικοί, είναι ιδιαίτερα ευαίσθητοι σε σφάλματα κατασκευής, ιδίως σε διακυμάνσεις του πλάτους του κυματοδηγού. Για παράδειγμα, μια μικρή διακύμανση του πλάτους κατά μόλις 5 nm μπορεί να προκαλέσει μετατόπιση συχνότητας κατά 250 GHz στη μετάδοση ενός μη ισορροπημένου MZI, κάτι που απαιτεί τη χρήση θερμικών ρυθμιστών υψηλής ισχύος για διόρθωση.

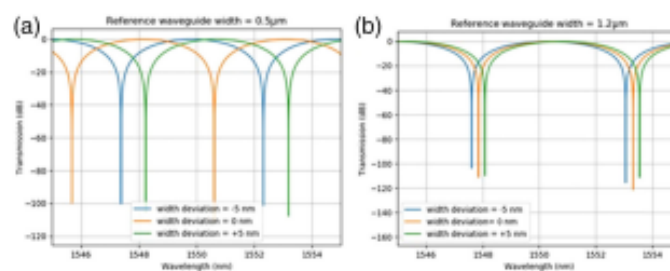
Στη δεύτερη φάση, θα αναπτυχθούν δομές WDM που θα είναι ανθεκτικές σε μεταβολές της κατασκευής έως και 2 GHz για διακύμανση πλάτους 5 nm. Τόσο τα αντηχεία δακτυλίου όσο και το καταρρακτικό MZI (βλέπε Σχ. 4.14) θα χρησιμοποιούν διαχωριστές φωτός με συσκευές πολλαπλών τρόπων λειτουργίας (MMI), οι οποίες είναι πιο ανθεκτικές στις

μεταβολές της κατασκευής. Επιπλέον, θα χρησιμοποιηθούν κυματοδηγοί με ευρύτερες καμπύλες που βασίζονται σε κάμψεις Όιλερ, οι οποίες αυξάνουν την ακτίνα καμπυλότητας για να αποτρέψουν τη διέγερση λειτουργιών υψηλότερης τάξης.

Στο Σχ. 4.15(α) παρουσιάζεται η μετάδοση ενός παραδοσιακού MZI με κυματοδηγό πλάτους 500 nm και στο Σχ. 4.15(β) φαίνεται ένα νέο σχέδιο MZI με κυματοδηγό πλάτους 1,2 μm. Στις δύο περιπτώσεις, αλλάζουμε το πλάτος του κυματοδηγού κατά ±5 nm και παρατηρούμε τις μετατοπίσεις στο μήκος κύματος μετάδοσης. Το νέο MZI με ευρύτερο κυματοδηγό εμφανίζει χαμηλότερη ευαισθησία σε αυτές τις αλλαγές, σε σύγκριση με τους τυπικούς κυματοδηγούς.



Σχήμα 4.14 Σχέδια κατασκευών MZI και WDM με μειωμένη ευαισθησία στις παραλλαγές κατασκευής.[5]



Σχήμα 4.15 Μετάδοση (α) ενός παραδοσιακά σχεδιασμένου MZI που βασίζεται στα 500 nm με κυματοδηγό απλής λειτουργίας και (β) MZI με βάση προκαταρκτικά σχέδια που αποτελούνται από το πρόσφατα σχεδιασμένο πλάτος 1,2 μm οδηγός κύματος.[5]

4.5 Αρχιτεκτονική Flex-SiPAC

Σε αυτή την ενότητα, παρουσιάζεται ο σχεδιασμός του Flex-SiPAC, δίνοντας έμφαση στις τεχνολογικές λεπτομέρειες των φωτονικών κυκλωμάτων πυριτίου (SiP) που χρησιμοποιούνται στην αρχιτεκτονική αυτή. Οι μαθηματικές εξισώσεις και τα σχετικά δεδομένα συνοψίζονται στον Πίνακα 4.1.

Notation	Description
r	Switch radix
l	Level index in the Flex-SiPAC topology, $l \in [0, L - 1]$, where $L = \max(l) + 1$
w	Number of tx/rx wavelengths per transceiver
N	Number of CUs in the topology
M	Number of WSSs in the topology
P	Number of jobs mapped to the topology
U	Set of CUs in job i : $U = \{u_i\}$, $i \in [0, P - 1]$
S	Set of WSSs in level l : $S = \{s_l\}$, $l \in [0, L - 1]$
T	Traffic matrix: $T = [t_{ij}] \in \mathbb{R}^{N \times N}$, $t_{ij} = 0$, $\forall i = j$
Λ	Number of wavelengths among each CU-pair (through l th level WSS): $\Lambda = [\lambda_{ij}^{(l)}] \in \mathbb{R}^{N \times N}$, $\lambda_{ij} = 0$, $\forall i = j$

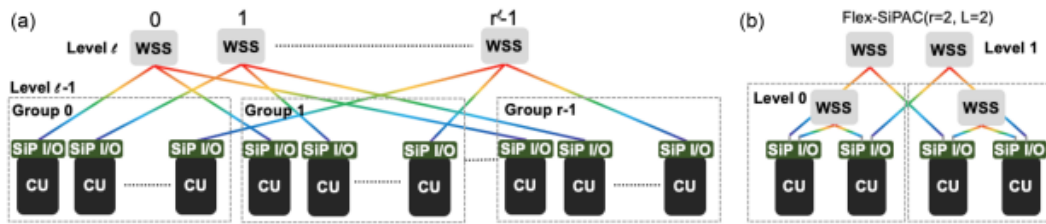
Πίνακας 4.1 Κατάλογος μαθηματικών συμβόλων με τις αντίστοιχες περιγραφές τους[6]

Σχεδιασμός τοπολογίας

Η αρχιτεκτονική Flex-SiPAC διατηρεί τον ίδιο βαθμό συνδεσιμότητας με την αρχιτεκτονική SiPAC, η οποία βασίζεται στην τοπολογία BCube. Το SiPAC(r, l) είναι μια αναδρομικά ορισμένη τοπολογία, όπου τα r και l αναπαριστούν τη ρίζα του switch και το τρέχον επίπεδο ή διάσταση στην τοπολογία. Η βασική μονάδα του επιπέδου $l = 0$ αποτελείται από έναν switch WSS με r θύρες, που συνδέει r κατανεμημένες υπολογιστικές μονάδες (CU).

Σύμφωνα με τη σχέση επανάληψης, για κάθε επίπεδο $l > 0$, το SiPAC(r, l) κατασκευάζεται με τη σύνδεση r διακοπών WSS, καθένας από τους οποίους συνδέει r μονάδες SiPAC από το προηγούμενο επίπεδο ($l - 1$), όπως φαίνεται στο Σχ. 4.16(α). Κάθε CU είναι εφοδιασμένη με $l + 1$ οπτικές θύρες, που αντιστοιχούν στον αριθμό των επιπέδων και τη διάμετρο της τοπολογίας.

Όπως και στο BCube, οι CU στο Flex-SiPAC λειτουργούν τόσο ως τερματικές μονάδες όσο και ως κόμβοι αναμετάδοσης. Για τη μείωση της κυκλοφοριακής συμφόρησης που προκαλείται από την αναμετάδοση, ο συλλογικός αλγόριθμος που θα παρουσιαστεί παρακάτω επιτρέπει σε κάθε CU να επικοινωνεί μόνο με τους άμεσους γείτονές της σε κάθε χρονικό βήμα με βάση έναν ενιαίο χάρτη εργασιών. Όταν χρειάζεται επικοινωνία μεταξύ απομακρυσμένων CU, ο αλγόριθμος εξασφαλίζει ότι το μήνυμα διασχίζει τα λιγότερα δυνατά ενδιάμεσα στάδια πριν ολοκληρωθεί η λειτουργία (π.χ., μείωση).



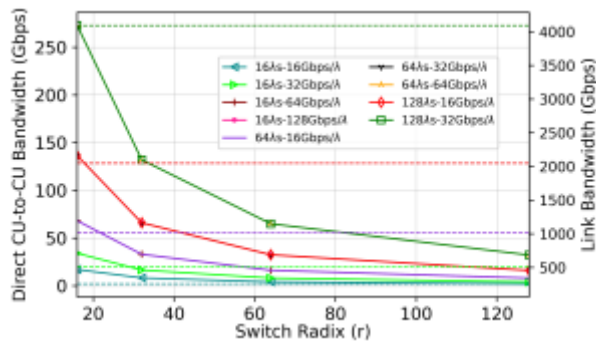
Σχήμα 4.16 α) Αρχιτεκτονική Flex-SiPAC που βασίζεται στην αναδρομική τοπολογία BCube . (β) Παράδειγμα δικτύου Flex-SiPAC($r = 2, L = 2$) που εμφανίζει δύο συνδέσεις θύρας ανά CU και πλήρη συνδεσιμότητα στο επίπεδο 0.[6]

Στην αρχιτεκτονική Flex-SiPAC, οι ηλεκτρονικοί switch(EPS) αντικαθίστανται από επαναδιαμορφώσιμα WSS (Wavelength Selective Switches), και η αρχιτεκτονική του πομποδέκτη που βασίζεται σε comb σε κάθε υπολογιστική μονάδα (CU) επανασχεδιάζεται για να ενσωματώσει τη δυνατότητα επαναδιαμόρφωσης στα τελικά σημεία. Αρχικά, τα μήκη κύματος από κάθε πομπό κατανέμονται ομοιόμορφα σε κάθε θύρα εξόδου, και το WSS σε κάθε διάσταση δρομολογεί τα μήκη κύματος ομοιόμορφα σε όλους τους άμεσα συνδεδεμένους γείτονες.

Είναι σημαντικό να σημειωθεί ότι οι υπάρχουσες πηγές λέιζερ comb προσφέρουν ήδη 160 μήκη κύματος , και οι φωτονικοί διαμορφωτές πυριτίου έχουν επιτύχει ρυθμούς δεδομένων 128 Gb/s ανά μήκος κύματος. Τα σχέδια κλιμακούμενων διακοπών αντηχείων μικροδακτυλίου (MRR) έχουν αποδείξει τη σκοπιμότητα για τον αριθμό θυρών έως 128. Ωστόσο, για την επίτευξη αυτών των αριθμών απαιτείται να αντιμετωπιστούν αρκετές προκλήσεις, όπως η διαχείριση της απώλειας εισαγωγής, οι διακυμάνσεις θερμοκρασίας, η ευθυγράμμιση του μηχανικού ελεύθερου φασματικού εύρους του MRR (FSR) με τα μήκη κύματος και η κλιμάκωση του αριθμού των θυρών μεταγωγής.

Το σχήμα 4.17 απεικονίζει την προβλεπόμενη τάση για μελλοντική κλιμάκωση εύρους ζώνης σε συνάρτηση με τη ρίζα του switch που χρησιμοποιείται στην αρχιτεκτονική Flex-SiPAC. Με μια ομοιόμορφη κατανομή μήκους κύματος, το εύρος ζώνης ανά ζεύξη μπορεί να φτάσει έως 4 Tb/s. Δεδομένου ότι κάθε CU σε ένα SiPAC(r, l) μπορεί να επιτύχει $(r - 1)$ άμεσους γείτονες μέσω δρομολόγησης μήκους κύματος σε μία μόνο διάσταση, το άμεσο εύρος ζώνης μεταξύ CU μπορεί να φτάσει έως 256 Gb/s όταν η ρίζα του switch είναι 16.

Η ρίζα WSS που χρησιμοποιείται είναι συνήθως μικρή, δεδομένου ότι ο αριθμός των τελικών σημείων κλιμακώνεται εκθετικά με τον αριθμό των επιπέδων στην τοπολογία. Για παράδειγμα, ένας switch radix 16 θα μπορούσε να επιτύχει μέγεθος τοπολογίας 256 για $L = 2$ και 4096 για $L = 3$. Επιπλέον, η δυνατότητα ευέλικτου συντονισμού των μηκών κύματος τόσο στον πομποδέκτη όσο και στον switch μας επιτρέπει να πετύχουμε έναν πολύ μεγαλύτερο σχεδιαστικό χώρο, ο οποίος χαρακτηρίζεται από το ολόκληρο διάστημα κάτω από τις διακεκομμένες οριζόντιες καμπύλες στο Σχ. 4.17 για ένα δεδομένο εύρος ζώνης σύνδεσης.



Σχήμα 4.17 Κλιμάκωση εύρους ζώνης Flex-SiPAC υπό ομοιόμορφη κατανομή των μηκών κύματος.[6]

Τεχνολογίες SiP για Flex-SiPAC

A) Switch με δυνατότητα επαναδιαμόρφωσης μήκους κύματος SiP

Το Wavelength Selective Switch (WSS) που προτείνεται στην αρχιτεκτονική SiPAC έχει τη δυνατότητα να δρομολογεί πολλαπλά μήκη κύματος μέσω μιας μόνο θύρας εισόδου προς κάθε έναν από τους άμεσα συνδεδεμένους γείτονές του, δημιουργώντας μια οπτική τοπολογία all-to-all με υψηλό εύρος ζώνης σε κάθε διάσταση. Αυτή η δυνατότητα ενισχύει την αποτελεσματική οπτική πολλαπλή μετάδοση, βελτιώνοντας την επικοινωνιακή αποδοτικότητα των συλλογικών λειτουργιών.

Σε αυτό το έργο, ενισχύεται αυτή τη σχεδίαση ενσωματώνοντας την χωρική δυνατότητα επαναδιαμόρφωσης στα WSS. Σκοπός είναι να δρομολογηθεί ένας ευέλικτος αριθμός εισερχόμενων μηκών κύματος από οποιαδήποτε θύρα εισόδου σε οποιαδήποτε θύρα εξόδου, αποφεύγοντας τις συνεκτικές διαφωνίες σε κάθε δίαυλο εξόδου.

Μια άλλη προσέγγιση περιλαμβάνει την αντικατάσταση των κανονικών αντηχειών με μικροδακτυλίους σε ένα cross-bar switch με ρυθμιζόμενα κύτταρα μεταγωγής, όπου τα ελεύθερα φασματικά εύρη (FSR) κάθε κελιού μεταγωγής μπορεί να προσαρμοστεί για να αποδεχθεί συγκεκριμένο αριθμό μηκών κύματος.

B) Πομποδέκτες με δυνατότητα επαναδιαμόρφωσης μήκους κύματος SiP

Η ενσωμάτωσή των θυρών οπτικών πομποδεκτών απευθείας στους διακομιστές chip interposer καταργεί την ανάγκη για ακριβές κάρτες διασύνδεσης (NIC). Στον συγκεκριμένο σχεδιασμό, κάθε υπολογιστική μονάδα (CU) στο FlexSiPAC είναι εξοπλισμένη με έναν ενσωματωμένο πομποδέκτη DWDM με comb. Επιπλέον, για την υποστήριξη μαζικά παράλληλης μεταφοράς δεδομένων DWDM, οι πομποδέκτες ενισχύονται με τη δυνατότητα αναδιαμόρφωσης μήκους κύματος, επιτρέποντας τη δυναμική δρομολόγηση μηκών κύματος σε διάφορες θύρες εξόδου.

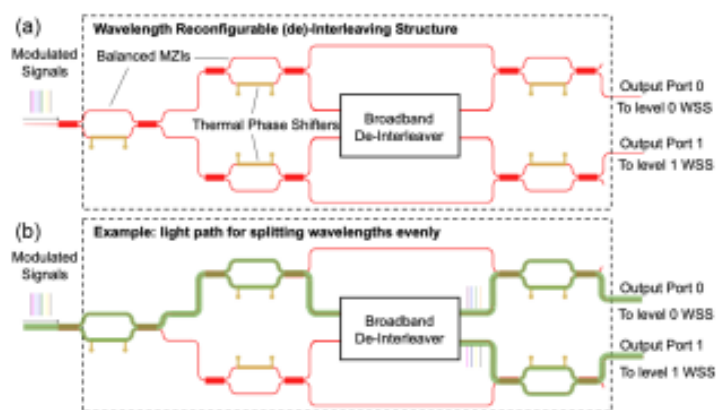
Κατά προεπιλογή, τα μήκη κύματος DWDM καταναμημένα ομοιόμορφα σε οπτικές ίνες που συνδέονται με τις CU σε διαφορετικές διαστάσεις. Ωστόσο, αν το μοτίβο κίνησης απαιτεί μεγαλύτερη χωρητικότητα από αυτή που μπορεί να προσφέρει μια μεμονωμένη διάσταση,

μπορούμε να ρυθμίσουμε τους πομποδέκτες ώστε να συνδυάζουν μήκη κύματος από διαφορετικές θύρες εξόδου και να τα μεταδίδουν μέσω μιας ενιαίας θύρας εξόδου.

Για την επίτευξη αυτού του στόχου, γίνεται ενσωμάτωση μιας αναδιαμορφώσιμης δομής αποσυμπίεσης μετά τον ενσωματωμένο πομποδέκτη, η οποία διαχωρίζει την αρχική έξοδο μιας ίνας σε πολλές εξόδους ινών. Το σχήμα 4.18(α) απεικονίζει ένα αρχιτεκτονικό παράδειγμα αυτού του σχεδιασμού δομής αποσυμπίεσης, όπου η κεντρική ευρυζωνική δομή αποσυμπίεσης υποστηρίζεται από αρκετά ισορροπημένα Mach-Zehnder συμβολόμετρα (MZI) για τη δρομολόγηση μηκών κύματος σε διαφορετικές διαδρομές. Σε αυτή τη διαμόρφωση πομποδέκτη 1×2 , τα DWDM μήκη κύματος που εισέρχονται στη δομή αποσυμπίεσης μπορούν να ρυθμιστούν ώστε να ταξιδεύουν μέσα από διάφορα μονοπάτια φωτός με τη βοήθεια των MZI. Για ομοιόμορφη κατανομή, τα μήκη κύματος δρομολογούνται μέσω του κέντρου ευρυζωνικής αποσυμπίεσης (Σχήμα 4.18(β)).

Η συγκεκριμένη υλοποίηση αυτής της δομής αποσυμπίεσης μπορεί να διαφοροποιείται ανάλογα με τις απαιτήσεις του συστήματος. Όταν απαιτείται να ανακατευθύνουμε τα μήκη κύματος από τη δεύτερη θύρα εξόδου προς την πρώτη, μπορούμε να ρυθμίσουμε τα MZI αναλόγως, ώστε όλα τα μήκη κύματος να δρομολογούνται μέσω του άνω κλάδου, παρακάμπτοντας τη δομή ευρυζωνικής αποσυμπίεσης. Συμμετρικές λειτουργίες μπορούν επίσης να εκτελούνται για τη δρομολόγηση όλων των μηκών κύματος στη δεύτερη θύρα εξόδου.

Αυτή η αναδιαμορφώσιμη δομή αποσυμπίεσης λειτουργεί ως βασική μονάδα που μπορεί να κλιμακωθεί σε πολλαπλά στάδια, δημιουργώντας ιεραρχικά περισσότερες αναδιαμορφώσιμες θύρες εξόδου. Ο αριθμός των απαιτούμενων θυρών είναι ίσος με τον αριθμό των επιπέδων στο FlexSiPAC, ο οποίος συνήθως είναι μικρός (π.χ. 2 ή 3). Η συνδυαστική επίδραση αυτού του συντονισμένου συνδυασμού πομποδέκτη-switch μας επιτρέπει να διαμορφώνουμε εκ νέου τα μήκη κύματος σε διαφορετικές διαστάσεις καθώς και μεταξύ διαφορετικών τελικών σημείων σε μια ενιαία διάσταση.



Σχήμα 4.18 (α) Παράδειγμα αρχιτεκτονικής πομποδέκτη με δυνατότητα επαναδιαμόρφωσης μήκους κύματος 1×2 . (β) Ομοιόμορφη κατανομή των εισερχόμενων διαμορφωμένων μήκη κύματος στις δύο θύρες εξόδου.[6]

Γ) Μηχανική Τοπολογία στο Wavelength Domain

Όσον αφορά τον αλγόριθμο διεύθυνσης εύρους ζώνης μήκους κύματος στο πλαίσιο του δικτύου Flex-SiPAC. Αρχίζουμε με την υπόθεση ότι οι εργασίες κατατάσσονται διαδοχικά στο σύμπλεγμα, όπου κάθε εργασία επιλέγει έναν συλλογικό αλγόριθμο, για την εκτέλεση συλλογικών λειτουργιών, βάσει παραλληλίας συμφερόντων.

Για παράδειγμα, όταν τέσσερις εργασίες μεγέθους 4 κατατάσσονται διαδοχικά σε ένα Flex-SiPAC ($r = 4, l = 1$) χρησιμοποιώντας τον αλγόριθμο πλέγματος, κάθε πίνακας υπο-επισκεψιμότητας σχηματίζει ένα πλήρες πλέγμα στις τέσσερις μονάδες ελέγχου (CU) κάτω από το ίδιο επίπεδο WSS 0. Αντίθετα, όταν δύο εργασίες μεγέθους 8 ανατίθενται, κάθε πίνακας δευτερεύουσας κυκλοφορίας δημιουργεί ένα πλήρες πλέγμα στις οκτώ CU, που εκτείνονται σε δύο επίπεδα WSS.

Δεδομένου του πίνακα υπο-επισκεψιμότητας που προκύπτει από τον επιλεγμένο συλλογικό αλγόριθμο για κάθε εργασία, συνθέτουμε έναν πλήρη πίνακα κυκλοφορίας τοπολογίας $T=[t_{ij}]$ συνδυάζοντας αυτούς τους πίνακες δευτερεύουσας κυκλοφορίας. Αυτοί οι αρχικοί πίνακες κυκλοφορίας, που χρησιμοποιούνται για τη βελτιστοποίηση της τοπολογίας, είναι απλοποιημένες αναπαραστάσεις των πραγματικών πινάκων κυκλοφορίας για την ανακατεύθυνση του εύρους ζώνης που αρχικά εκτείνεται σε CU, σε πολλές εργασίες έως και εντός CU που ανήκουν στην ίδια εργασία.

Στην περίπτωση του Flex-SiPAC, επιλέγουμε το SiPCO για να περιορίσουμε την επικοινωνία στους άμεσα συνδεδεμένους γείτονες. Αυτό διασφαλίζει ότι $t_{ij}, \forall i, j: l_{ij}=0$. Στη συνέχεια, απορρίπτουμε κάθε κίνηση που διασχίζει CU μεταξύ διαφορετικών εργασιών. Ο στόχος μας είναι διπλός: 1) να προσδιορίσουμε πώς τα μήκη κύματος από τη θύρα εισόδου του πομποδέκτη θα πρέπει να ανατεθούν σε κάθε θύρα εξόδου και 2) να διαμορφώσουμε τις καταστάσεις του switch σε κάθε διάσταση για να μεγιστοποιήσουμε την απόδοση του δικτύου.

Περιγράφουμε την επιθυμητή τοπολογία στον τομέα μήκους κύματος ως $\Lambda=[\lambda_{ij}]$. Για να ευθυγραμμίσουμε την τοπολογία με τον πίνακα κυκλοφορίας, στοχεύουμε να αναθέσουμε μήκη κύματος μεταξύ των CU σε αναλογία με τον όγκο της κίνησης που ανταλλάσσεται μεταξύ τους. Για να το επιτύχουμε αυτό, χρησιμοποιούμε μια συνταγή ILP όπως εκφράζεται στην Εξίσωση (1).

Η μεταβλητή απόφασης λ_{ij} αντιπροσωπεύει τον αριθμό των μηκών κύματος που συνδέουν το Cu_i με το Cu_j . Ο στόχος του προβλήματος βελτιστοποίησης περιλαμβάνει μια τετραγωνική αντικειμενική συνάρτηση που ελαχιστοποιεί τη διαφορά ελάχιστου τετραγώνου μεταξύ της τοπολογίας μήκους κύματος και του πίνακα κυκλοφορίας στόχου. Η επιλογή μιας τετραγωνικής αντικειμενικής συνάρτησης καθοδηγείται από την κυρτή της φύση, η οποία διευκολύνει τη σύγκλιση του αλγορίθμου βελτιστοποίησης.

Αυτή η βελτιστοποίηση υπόκειται σε 1) περιορισμούς εισόδου και 2) περιορισμούς εξόδου, που διασφαλίζουν ότι ο αριθμός των μηκών κύματος παραμένει εντός των ορίων των διαθέσιμων μηκών κύματος πομπού/δέκτη. Δεδομένου ότι αυτή η βελτιστοποίηση θεωρεί το συνολικά διαθέσιμο εύρος ζώνης σε όλες τις διαστάσεις, η λύση στο πρόβλημα ILP αντιστοιχεί στον συνολικό αριθμό των μηκών κύματος ανά ζεύγος CU.

Για να μειώσουμε την πολυπλοκότητα, χαλαρώνουμε τον τελευταίο περιορισμό από έναν ακέραιο σε μια συνεχή μεταβλητή απόφασης και χρησιμοποιούμε μια ευρετική προσέγγιση για να αποκτήσουμε μια λύση που προσεγγίζει τη βέλτιστη λύση. Η ευρετική περιλαμβάνει δύο βήματα: τη σωστή στρογγυλοποίηση των μεταβλητών απόφασης και μια επαναληπτική διαδικασία ανάθεσης για τα CU που δεν έχουν εκχωρηθεί το πλήρες σύνολο των μηκών κύματος w . Στο πρώτο βήμα, στρογγυλοποιούμε τις μεταβλητές απόφασης στις πλησιέστερες ακέραιες τιμές τους. Στη συνέχεια, εξετάζουμε κάθε CU που απαιτεί επιπλέον μήκη κύματος, εκχωρώντας βάρη με βάση τη διαφορά μεταξύ των επί του παρόντος κατανομημένων μηκών κύματος και του ιδανικού αριθμού μηκών κύματος. Τέλος, τα μήκη κύματος κατανέμονται ένα κάθε φορά με στρογγυλοποίηση που βασίζεται στη μείωση των τιμών βάρους.

Είναι σημαντικό να επισημάνουμε ότι η εκχώρηση μήκους κύματος σε κάθε θύρα του πομποδέκτη και στο WSS δεν είναι μια αμελητέα διαδικασία. Σε κάθε βήμα αυτής της διαδικασίας ανάθεσης μήκους κύματος, είναι απαραίτητο να διασφαλίσουμε ότι πληρούνται οι περιορισμοί εισόδου και εξόδου μήκους κύματος, καθώς και οι περιορισμοί πολυπλεξίας μήκους κύματος σε κάθε CU, σε ολόκληρη την τοπολογία. Αυτή η επαναληπτική διαδικασία συνεχίζεται μέχρι να εκχωρηθεί σε κάθε CU το πλήρες σύνολο των μήκων κύματος εκπομπής w .

Η δρομολόγηση μήκους κύματος σε κάθε WSS μπορεί να προσεγγιστεί ως ένα πρόβλημα χρωματισμού κορυφής, όπου τα επιμέρους σήματα αναπαρίστανται ως κόμβοι που συνδέονται με βάση διακριτές ιδιότητες, όπως είσοδος, έξοδος και μήκος κύματος. Αλγόριθμοι όπως ο DSatur θα μπορούσαν να χρησιμοποιηθούν για την επίλυση αυτού του προβλήματος:

$$\begin{aligned} & \min_{\Lambda} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (\lambda_{ij} - t_{ij})^2, \\ \text{s.t. } & 1) \sum_{j=1}^N \lambda_{ij} \leq w, \quad \forall 0 \leq i \leq N-1, \\ & 2) \sum_{i=1}^N \lambda_{ij} \leq w, \quad \forall 0 \leq j \leq N-1, \\ & 3) \lambda_{ii} = 0, \lambda_{ij} \in \mathbb{Z}, \quad \forall i, j \in \{0, \dots, N-1\}. \quad \mathbf{(1)} \end{aligned}$$

Για να μετατρέψουμε την ευρετική λύση σε έλεγχο δικτύου, μπορούμε να διαιρέσουμε τη διαδικασία σε δύο κύρια στοιχεία: τη ρύθμιση του πομποδέκτη και τη ρύθμιση του switch. Αυτό απαιτεί έναν κεντρικό ελεγκτή δικτύου με δυνατότητες SDN, ο οποίος είναι υπεύθυνος για τον συντονισμό των ενημερώσεων στους πομποδέκτες και τους switch.

Ο ελεγκτής δικτύου έχει επίσης την ευθύνη της συλλογής πληροφοριών σχετικά με τη χαρτογράφηση των θέσεων εργασίας στην τοπολογία και την επιλογή συλλογικών αλγορίθμων ανά εργασία. Σε κάθε χρονικό βήμα αναδιαμόρφωσης t , ο ελεγκτής εκτελεί μια διαδικασία ευρετικής βελτιστοποίησης για να καθορίσει τη βέλτιστη λύση για την κατανομή εύρους ζώνης $\lambda_{ij}[t]$.

Επαναλαμβάνουμε αυτή τη διαδικασία για όλα τα ζεύγη CU και διαπιστώνουμε αν το τρέχον σχέδιο κατανομής μήκους κύματος διαφέρει από αυτό του προηγούμενου χρονικού βήματος $l_{ij}[t-1]$. Αν υπάρχει διαφορά, καθορίζουμε τον αριθμό των μηκών κύματος που χρειάζονται αλλαγή και ρυθμίζουμε τόσο τον πομποδέκτη όσο και το WSS στις επιθυμητές καταστάσεις. Στόχος μας είναι να ελαχιστοποιήσουμε τις τροποποιήσεις στην υπάρχουσα κατανομή μήκους κύματος, εκχωρώντας μήκη κύματος που δεν έχουν ακόμα δοθεί σε έναν προορισμό στην τρέχουσα κατανομή.

Είναι σημαντικό να σημειωθεί ότι υποθέτουμε ότι η συχνότητα επαναδιαμόρφωσης βασίζεται σε μια φορά ανά φόρτο εργασίας, καθώς το μοτίβο κυκλοφορίας παραμένει σχετικά σταθερό σε κάθε εκπαιδευτική διαδικασία.

4.6 Συμμετοχικός Σχεδιασμός Συλλογικού Αλγορίθμου

Ο συλλογικός αλγόριθμος Flex-SiPCO αποτελεί μια επέκταση του αλγορίθμου SiPCO [18] και ενσωματώνει τη χαρτογράφηση εργασιών με την αναδιαμόρφωση εύρους ζώνης. Ακολουθεί μια παρόμοια σχεδιαστική προσέγγιση με τον αλγόριθμο SiPCO, εκμεταλλευόμενος τα πλεονεκτήματα της δυνατότητας πολλαπλών θυρών του Flex-SiPAC και αξιοποιώντας όλα τα διαθέσιμα μήκη κύματος σε κάθε διάσταση σε κάθε βήμα της διαδικασίας.

Παρακάτω παρουσιάζεται το Flex-SiPCO all-reduce ως ένα παράδειγμα, σημειώνοντας ότι μπορούν να προκύψουν και άλλες συλλογικές λειτουργίες, όπως το all-gather και το reduce-scatter, από το all-reduce, και ότι μια παρόμοια αρχή σχεδιασμού μπορεί να εφαρμοστεί στην επικοινωνία όλων προς όλους.

Algorithm 1. Flex-SiPCO All-Reduce Algorithm

Input: $r, U = \{u_i | i \in [0, P-1]\}$, $\Lambda = [\lambda_{ij}] \in \mathbb{R}^{N \times N}$

- 1: **for** each job $i \in [0, P-1]$ **do**
- 2: $u_i \leftarrow$ Set of CUs in job i
- 3: $\hat{l} \leftarrow \max\{l | \text{removing } s_l \text{ disconnects } u_i\} + 1$
- 4: **for** each $CU_x \in u_i, i \in [0, P-1]$ **do**
- 5: Partition the local message into $r\hat{l}$ evenly sized chunks
- 6: $C_x^c \leftarrow$ chunk $c \in [0, r-1]$ in group $g \in [0, \hat{l}-1]$
- 7: \triangleright Label each r contiguous chunks with index $g \in [0, \hat{l}-1]$
- 8: \triangleright Label each chunk within each group $c \in [0, r-1]$
- 9: **for** each $l \in [0, \hat{l}-1]$ **do**
- 10: **for** each $y \in [0, r-1] | \lambda_{xy}^l > 0$ **do**
- 11: Send $\{C_x^c\}$ using link $l \rightarrow CU_y, \forall CU_y \in u_i$
- 12: Send $\{C_x^c\}$ using link $(l+1) \bmod \hat{l}, \forall CU_y \notin u_i$
- 13: **for** step $v \in [1, \hat{l}]$ **do**
- 14: **for** each $CU_x \in u_i, i \in [0, P-1]$ **do**
- 15: **for** each $l \in [0, \hat{l}-1]$ **do**
- 16: $g = (v+l) \bmod \hat{l}$
- 17: $rc =$ reduced chunks, $fc =$ transit chunks
- 18: Bcast $\{C_x^g\}$ using link $l \rightarrow CU_y, \forall CU_y \in u_i,$
 $\lambda_{xy}^l > 0$
- 19: Forward $\{C_x^g\}$ to level $(l+1) \bmod \hat{l}$

Το κόστος καθυστέρησης του αλγορίθμου Flex-SiPCO μπορεί να αναπαρασταθεί με το μοντέλο $(a + nb)$ όπου:

a είναι η καθυστέρηση σύνδεσης ανά χρονικό βήμα.

b είναι ο χρόνος μεταφοράς ανά byte.

n είναι το μέγεθος του μηνύματος σε έναν σύνδεσμο ανά βήμα μονάδας.

Αυτό το μοντέλο καταδεικνύει ότι η συνολική καθυστέρηση είναι το άθροισμα της σταθερής καθυστέρησης σύνδεσης και του δυναμικού χρόνου μεταφοράς που εξαρτάται από το μέγεθος του μηνύματος. Συνοψίζοντας, το Flex-SiPCO αποτελεί μια γενίκευση του προηγούμενου προτεινόμενου αλγορίθμου SiPCO, καθώς ενσωματώνει τη χαρτογράφηση εργασιών και την αναδιαμόρφωση της τοπολογίας. Όταν η κατανομή εργασίας ευθυγραμμίζεται με το μέγεθος της τοπολογίας, ο αλγόριθμος επιστρέφει στην αρχική μορφή του SiPCO.

Συμπεράσματα

Οι σύγχρονες τάσεις στην οπτική δικτύωση επικεντρώνονται στην εξυπηρέτηση εφαρμογών που απαιτούν υψηλή απόδοση στα data centers, με σκοπό τη βελτίωση της αποδοτικότητας, της ταχύτητας και της ασφάλειας των μεταφορών δεδομένων. Η χρήση οπτικών ινών προσφέρει τη δυνατότητα μεταφοράς μεγάλων όγκων δεδομένων με ελάχιστη καθυστέρηση και χαμηλή κατανάλωση ενέργειας, καθιστώντας τις ιδανικές για απαιτητικές εφαρμογές όπως το cloud computing, η διαχείριση μεγάλων δεδομένων (big data) και οι υπηρεσίες τεχνητής νοημοσύνης. Κάποια από τα βασικά χαρακτηριστικά αυτών των σύγχρονων τάσεων είναι: 1) η ταχύτητα και απόδοση, οι οπτικές ίνες προσφέρουν εξαιρετικά γρήγορες μεταφορές δεδομένων, απαραίτητες για εφαρμογές που απαιτούν άμεση πρόσβαση σε πληροφορίες, 2) η όλο και μεγαλύτερη χωρητικότητα, η τεχνολογία Dense Wavelength Division Multiplexing (DWDM) επιτρέπει τη συγχώνευση πολλών σημάτων σε μία μόνο ίνα, μεγιστοποιώντας τη χωρητικότητα και επιτρέποντας τη μεταφορά πολλαπλών ροών δεδομένων ταυτόχρονα, 3) η ασφάλεια των δεδομένων, οι οπτικές επικοινωνίες προσφέρουν καλύτερη ασφάλεια σε σύγκριση με τις ηλεκτρικές συνδέσεις, καθώς είναι λιγότερο επιρρεπείς σε παρεμβολές και επιθέσεις και 4) οι προοπτικές ευελιξίας και επεκτασιμότητας, η ευελιξία και η επεκτασιμότητα είναι κρίσιμες για την υποστήριξη της συνεχώς αυξανόμενης ζήτησης για ταχύτερες και αξιόπιστες συνδέσεις. Τέλος, οι εξελίξεις στην οπτική δικτύωση είναι καθοριστικές για την ανάπτυξη των σύγχρονων υποδομών πληροφορικής, καθώς συμβάλλουν στην αποτελεσματική διαχείριση των δεδομένων και την ικανοποίηση των ολοένα και αυξανόμενων αναγκών των χρηστών.

Βιβλιογραφία

- [1] W. Xia, P. Zhao, Y. Wen and H. Xie, "A Survey on Data Center Networking (DCN): Infrastructure and Operations," in *IEEE Communications Surveys & Tutorials*, vol. 19, no. 1, pp. 640-656
- [2] Sankaran, G.C., Sivalingam, K.M. A survey of hybrid optical data center network architectures. *Photon Netw Commun* **33**, 87–101 (2017). <https://doi.org/10.1007/s11107-016-0643-2>
- [3] Algorithms for Intelligent Systems Series Editors: Jagdish Chand Bansal · Kusum Deep · Atulya K. Nagar , V. K. Giri Nishchal K. Verma R. K. Patel V. P. Singh Editors
- [4] Zacharaya SHABKA, Optimisation for Optical Data Centre Switching and Networking with Artificial Intelligence
- [5] Madeleine Glick, Nathan C. Abrams, Qixiang Cheng, Min Yee Teh, Yu-Han Hung, Oscar Jimenez, Songtao Liu, Yoshitomo Okawachi, Xiang Meng, Leif Johansson, Manya Ghobadi, Larry Dennison, George Michelogiannakis, John Shalf, Alan Liu, John Bowers, Alex Gaeta, Michal Lipson, and Keren Bergman, "PINE: Photonic Integrated Networked Energy efficient datacenters (ENLITENED Program) [Invited]," *J. Opt. Commun. Netw.* 12, 443-456 (2020)
- [6] Zhenguo Wu, Liang Yuan Dai, Yuyang Wang, Songli Wang, and Keren Bergman, "Flexible silicon photonic architecture for accelerating distributed deep learning," *J. Opt. Commun. Netw.* 16, A157-A168 (2024)
- [7] Tech Target Definitions. <https://www.techtarget.com>

