

- 1) Ο ορισμός που δίνεται από τους συγγραφείς είναι: «Τα p2p συστήματα είναι μια κλάση συστημάτων και εφαρμογών που χρησιμοποιούν καταναμημένους πόρους για την εκτέλεση μιας συνάρτησης με μη κεντρικοποιημένο τρόπο». Ο ορισμός που έχει δώσει η ομάδα της Intel που ασχολείται με P2P δίκτυα είναι: « Ένα P2P σύστημα είναι ο διαμοιρασμός υπολογιστικών πόρων και υπηρεσιών με άμεση ανταλλαγή μεταξύ των συστημάτων». Οι δύο αυτοί ορισμοί μπορούν να θεωρηθούν σχεδόν ταυτόσημοι. Η διαφορά είναι ότι στον δεύτερο τονίζεται η δυνατότητα διαμοιρασμού των πόρων που υπάρχουν σε κάθε peer, ενώ στον πρώτο τονίζεται η δυνατότητα χρήσης των πόρων που προσφέρονται από διάφορα peer για την εκτέλεση μιας εργασίας το αποτέλεσμα της οποίας θα προκύψει με συνεργασία όλων αυτών των κόμβων.
- 2) Client-server model vs p2p model
 - managed vs self-organized: Στο client-server μοντέλο υπάρχει ένας υπολογιστής που διαχειρίζεται όλες τις αιτήσεις, ενώ στο p2p μοντέλο, όλοι οι κόμβοι συμμετέχουν στην επιλογή του κόμβου που θα εξυπηρετήσει την αίτηση.
 - Configured vs ad-hoc: Στο client-server μοντέλο οι συνδέσεις μεταξύ των κόμβων είναι καθορισμένες και το δίκτυο δε μεταβάλλεται συχνά, ενώ στο p2p μοντέλο οι αλλαγές είναι αρκετά συχνές.
 - Lookup vs discover: Στο client-server μοντέλο είναι γνωστός ο κόμβος που προσφέρει την επιθυμητή υπηρεσία και αρκεί η αναζήτηση της υπηρεσίας στον κόμβο αυτό, ενώ στο p2p μοντέλο ο κόμβος που προσφέρει την υπηρεσία δεν είναι γνωστός και χρειάζεται εξερεύνηση του δικτύου για να βρεθεί.
 - Hierarchy vs mesh: Στο client-server μοντέλο υπάρχει μια ιεραρχία έτσι ώστε κάθε κόμβος να απευθύνεται στον αμέσως παραπάνω στην ιεραρχία, ενώ στο p2p μοντέλο δεν είναι καθορισμένος ο κόμβος στον οποίο πρέπει να γίνει αναφορά αλλά καθορίζεται από την ερώτηση και από το μηχανισμό δρομολόγησης κάθε κόμβου.
 - Static vs mobile: Στο client-server μοντέλο η τοπολογία του δικτύου δεν μεταβάλλεται, ενώ στο p2p μοντέλο μπορεί να αλλάξει αν για παράδειγμα κάποιος κόμβος αποσυνδεθεί και συνδεθεί σε διαφορετική τοποθεσία
 - Server dependencies vs independent lifetime: Στο client-server μοντέλο αν ο server για κάποιο λόγο πέσει το μοντέλο δεν μπορεί να λειτουργήσει και οι χρήστες δεν μπορούν να εξυπηρετηθούν, ενώ στο p2p μοντέλο ακόμη και αν κάποιος κόμβος τεθεί εκτός λειτουργίας το δίκτυο συνεχίζει να λειτουργεί κανονικά με τη διαφορά ότι ο συγκεκριμένος κόμβος δεν μπορεί να προσφέρει τις υπηρεσίες του.
 - DNS-based vs custom naming: Στο client-server μοντέλο το όνομα κάθε κόμβου καθορίζεται και αποθηκεύεται στον DNS server, ενώ στο p2p μοντέλο κάθε κόμβος μπαίνει στο σύστημα με το όνομα που θέλει και το οποίο το μεταδίδει στους γείτονες του.
 - RPC vs asynchronous: Στο client-server μοντέλο χρησιμοποιείται σύγχρονη επικοινωνία με κλήση απομακρυσμένων διαδικασιών, ενώ στο p2p μοντέλο η επικοινωνία είναι ασύγχρονη αφού οι αλλαγές στο δίκτυο είναι πολύ συχνές και οι χρήστες περιοδικά μπαίνουν και βγαίνουν από το σύστημα.
- 3) Ένα client-server μοντέλο είναι *flat* αν όλοι οι clients επικοινωνούν μόνο με έναν μοναδικό server. Αντίθετα είναι *hierarchical* αν οι servers ενός επιπέδου λειτουργούν ως clients για τους κόμβους υψηλότερων επιπέδων.

- 4) Οι super peers είναι κάποιιοι κόμβοι ενός P2P δικτύου οι οποίοι διατηρούν κάποιες πληροφορίες που άλλοι κόμβοι μπορεί να μην τις έχουν. Οι άλλοι κόμβοι ρωτάνε τους super peers για πληροφορίες τις οποίες δεν μπορούν να βρουν με άλλο τρόπο.
- 5) Η διαφορά ανάμεσα στα δύο αυτά είδη εφαρμογών είναι ότι στις compute-intensive εφαρμογές η παραλληλοποίηση γίνεται εκτελώντας την ίδια εργασία σε κάθε peer αλλά με διαφορετικές παραμέτρους, ενώ στις componentized εφαρμογές η παραλληλοποίηση γίνεται εκτελώντας μια διαφορετική εργασία σε κάθε peer. Ο πρώτος τρόπος παραλληλοποίησης ταιριάζει περισσότερο σε περιπτώσεις που χρησιμοποιείται οριζόντια κατανομή για την αποθήκευση των δεδομένων, ενώ ο δεύτερος τρόπος ταιριάζει σε περιπτώσεις που χρησιμοποιείται κατακόρυφη κατανομή. Για παράδειγμα, ας θεωρήσουμε την περίπτωση μιας κατανεμημένης βάσης δεδομένων. Αν σε κάθε κόμβο αποθηκεύεται ένας αριθμός πλειάδων του πίνακα (οριζόντια κατανομή), τότε για την απάντηση μιας ερώτησης αρκεί να εφαρμοστεί η ίδια ερώτηση σε κάθε κόμβο (compute-intensive). Αντίθετα αν αποθηκεύεται σε κάθε κόμβο αποθηκεύεται ένα υποσύνολο των γνωρισμάτων του πίνακα (κατακόρυφη κατανομή), τότε πρέπει να υποβληθεί μια διαφορετική ερώτηση σε κάθε κόμβο (componentized).
- 6) Η κύρια πρόκληση όσον αφορά την επικοινωνία σε P2P δίκτυα είναι η αντιμετώπιση των προβλημάτων που σχετίζονται με τη δυναμική φύση των peers. Το σύνολο των peers αλλάζει συχνά είτε επειδή ένας χρήστης κλείνει τον υπολογιστή του, είτε από σφάλματα στο δίκτυο. Η διατήρηση συνδεσιμότητας σε επίπεδο εφαρμογής σε ένα τέτοιο δίκτυο είναι μία από τις μεγαλύτερες προκλήσεις στην ανάπτυξη P2P εφαρμογών.
- 7) Ο βασικότερος τρόπος αντιμετώπισης των προβλημάτων αξιοπιστίας που παρουσιάζουν τα P2P δίκτυα είναι η χρήση επανάληψης. Δηλαδή τα ίδια δεδομένα και υπηρεσίες υπάρχουν σε περισσότερους από ένα κόμβους ώστε να μπορεί να εξυπηρετηθεί ο χρήστης ακόμη και αν κάποιος από αυτούς τους κόμβους είναι προσωρινά εκτός λειτουργίας ή συμβεί κάποιο σφάλμα στην προσπάθεια επικοινωνίας με κάποιο κόμβο.
- 8) Τα πλεονεκτήματα και μειονεκτήματα του αλγορίθμου που χρησιμοποιεί centralized directory είναι:

Πλεονεκτήματα:

- Επιλογή του καλύτερου peer που ταιριάζει με την ερώτηση
- Αφού βρεθεί το peer η ανταλλαγή αρχείων γίνεται άμεσα μεταξύ των δύο peers
- Δεν επιβαρύνεται το δίκτυο με πολλά μηνύματα

Μειονεκτήματα:

- Ο directory server πρέπει να κρατά πληροφορία για όλους τους κόμβους του δικτύου
- Ο directory server μπορεί να αποτελέσει σημείο συμφόρησης
- Τα δύο παραπάνω περιορίζουν τις δυνατότητες κλιμάκωσης

Όσον αφορά το flooded requests model έχουμε:

Πλεονεκτήματα:

- Οι κόμβοι δε χρειάζεται να κρατούν πληροφορία για άλλους κόμβους του δικτύου

- Σίγουρη εύρεση απάντησης (αν αυτή υπάρχει μέσα στο όριο του μέγιστου αριθμού βημάτων που καθορίζεται από τον αλγόριθμο)

Μειονεκτήματα:

- Μεγάλος αριθμός μηνυμάτων, απαίτηση για μεγάλο bandwidth, προβλήματα κλιμάκωσης
- Περιττά μηνύματα σε κόμβους που δεν μπορούν να εξυπηρετήσουν την ερώτησή μας

Τέλος για το document routing model έχουμε:

Πλεονεκτήματα:

- Με χρήση της συνάρτησης κατακερματισμού μπορεί να βρεθεί ο κόμβος στον οποίο είναι αποθηκευμένο ένα αρχείο

Μειονεκτήματα:

- Το ID ενός αρχείου πρέπει να είναι γνωστό πριν γίνει κάποια αίτηση για αυτό
- Κάθε κόμβος πρέπει να έχει τις απαραίτητες πληροφορίες για την δρομολόγηση κάθε αίτησης στον κατάλληλο κόμβο.

- 9) Το πλεονέκτημα είναι ότι αρκούν δύο μηνύματα για την ανταλλαγή του αρχείου. Δηλαδή δεν χρειάζεται να γίνει δρομολόγηση της αίτησης μέσω άλλων κόμβων. Από την άλλη πλευρά υπάρχει το μειονέκτημα της ύπαρξης ενός μόνο κόμβου επικοινωνίας, ο οποίος αν για κάποιο λόγο τεθεί εκτός λειτουργίας τότε δεν μπορεί να γίνει η ανταλλαγή του αρχείου.
- 10) Στο Gnutella οι νέοι κόμβοι πρέπει να ξέρουν την διεύθυνση κάποιου άλλου κόμβου ή να χρησιμοποιήσουν κάποια λίστα με γνωστές διευθύνσεις άλλων κόμβων. Ο νέος κόμβος εισέρχεται στο δίκτυο εγκαθιστώντας σύνδεση προς έναν τουλάχιστον από τους υπάρχοντες κόμβους. Μετά μπορεί να αρχίσει την εξερεύνηση άλλων κόμβων και να αποθηκεύει τοπικά τις IP διευθύνσεις τους.
- 11) Οι παράγοντες που επηρεάζουν την κλιμάκωση είναι: α) Ο αριθμός των κεντροποιημένων λειτουργιών. Για παράδειγμα αν όλοι οι κόμβοι του δικτύου ζητούν κάποια υπηρεσία από έναν μόνο υπολογιστή τότε περιορίζεται η δυνατότητα κλιμάκωσης. β) Η ποσότητα των καταστάσεων που πρέπει να διατηρούνται (π.χ. αν απαιτείται η αποθήκευση μεγάλης ποσότητας πληροφορίας από κάποιους κόμβους). γ) Ο έμφυτος παραλληλισμός που παρουσιάζει η εφαρμογή. Για παράδειγμα μπορεί μια εφαρμογή λόγω της φύσης της να μην μπορεί να χωριστεί σε κομμάτια που καθένα από αυτά να εκτελεστούν σε διαφορετική μηχανή ώστε να επιταχυνθεί η εκτέλεση. δ) Το προγραμματιστικό μοντέλο που χρησιμοποιείται για τους υπολογισμούς, το οποίο μπορεί να επιτρέπει ή όχι την παράλληλη εκτέλεση και τον συνδυασμό αποτελεσμάτων από διαφορετικούς κόμβους. Επίσης η ταχύτητα της γλώσσας προγραμματισμού μπορεί να παίξει καθοριστικό ρόλο καθώς και το μοντέλο που μπορεί να χρησιμοποιηθεί για κλήση διαδικασιών σε απομακρυσμένους υπολογιστές (π.χ. RPC και RMI).
- 12) Ο καταλληλότερος τρόπος επικοινωνίας είναι ο transient asynchronous αφού οι αλλαγές στο δίκτυο είναι πολύ συχνές και οι χρήστες περιοδικά μπαίνουν και βγαίνουν από το σύστημα. Επίσης είναι συχνά τα σφάλματα στο δίκτυο λόγω του μεγάλου αριθμού βημάτων που μπορεί να διανύσει κάθε μήνυμα.
- 13) Αυτό που εννοούν οι συγγραφείς με αυτή τη φράση είναι ότι όσο περισσότερο διασπαρμένοι σε διάφορα σημεία είναι οι κόμβοι ενός p2p δικτύου, τόσο καλύτερη είναι η απόδοση λόγω του περιορισμού στο φόρτο

του δικτύου και των κόμβων. Αν δηλαδή υπάρχουν πολλά peers με τα ίδια δεδομένα είναι προτιμότερο οι κόμβοι αυτοί να βρίσκονται σε διάφορα γεωγραφικά σημεία, έτσι ώστε αν κάποιος χρήστης ζητήσει αυτά τα δεδομένα να υπάρχει κοντά του κάποιος κόμβος που να μπορεί να τον εξυπηρετήσει. Έτσι περιορίζεται ο φόρτος του δικτύου αφού τα μηνύματα δεν διανύουν μεγάλες αποστάσεις σε αυτό, καθώς και ο φόρτος των peer γιατί καθένα θα δέχεται αιτήσεις μόνο από τους κόμβους οι οποίοι είναι κοντά του στο δίκτυο.

- 14) Στόχος είναι η ελαχιστοποίηση του χρόνου πρόσβασης στα peers, η ταχύτερη απόκριση στις ερωτήσεις και η ισοκατανομή του φόρτου στο σύστημα. Η χρήση cache για την αποθήκευση των αποτελεσμάτων σε όλους τους κόμβους του μονοπατιού έχει το πλεονέκτημα ότι οποιαδήποτε ερώτηση φτάσει σε κάποιον από αυτούς τους κόμβους μπορεί να απαντηθεί από τα δεδομένα τις cache χωρίς να είναι απαραίτητος ο εντοπισμός των κόμβων από τους οποίους προέρχονται τα δεδομένα, μειώνοντας έτσι την καθυστέρηση απόκρισης και τον αριθμό των βημάτων που διανύει κάθε μήνυμα στο δίκτυο. Ένα βασικό μειονέκτημα είναι το κόστος ενημέρωσης, αφού για κάθε αλλαγή στα δεδομένα πρέπει να ενημερωθούν οι κόμβοι οι οποίοι τα διατηρούν στην cache τους. Επίσης υπάρχει σπατάλη αποθηκευτικού χώρου σε κάθε peer και επιβάρυνσή του με την αποθήκευση και ενημέρωση των δεδομένων. Αυτό μπορεί να γίνει και στο Gnutella (στην οποία ακολουθείται το flooded requests model) αν κρατάμε το μονοπάτι που ακολούθησε το μήνυμα για να φτάσει στον κόμβο με την απάντηση. Ένας εναλλακτικός τρόπος με μικρότερο κόστος ενημέρωσης και αποθήκευσης είναι η αποθήκευση του αποτελέσματος μόνο στην cache του peer που έκανε την ερώτηση.
- 15) Σημαίνει ότι υπάρχει ένας αριθμός από κόμβους με πολλές συνδέσεις (μεγάλο βαθμό), οι οποίοι βοηθούν στην επιτάχυνση της αναζήτησης με την προώθηση των αιτήσεων προς αυτούς τους κόμβους.
- 17) Το *fault tolerance* είναι μεγαλύτερη πρόκληση σε collaborative p2p systems από ότι σε sharing p2p systems για το λόγο ότι στην πρώτη περίπτωση η παράδοση των μηνυμάτων πολλές φορές πρέπει να είναι αξιόπιστη και επίσης πολλές φορές η σειρά με την οποία έρχονται τα μηνύματα είναι σημαντική. Όλα αυτά κάνουν πιο δύσκολη τη σωστή λειτουργία του συστήματος και την ανάκαμψη από σφάλματα.