

**ΟΙΚΟΝΟΜΙΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ**



**ATHENS UNIVERSITY
OF ECONOMICS
AND BUSINESS**

Κινητά και Διάχυτα Συστήματα

Ενότητα # 6: Εφαρμογές DHT
Διδάσκων: Γεώργιος Ξυλωμένος
Τμήμα: Πληροφορικής



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Οικονομικό Πανεπιστήμιο Αθηνών**» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΣΠΑ
2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Άδειες Χρήσης

- Το παρόν εκπαιδευτικό υλικό υπόκειται σε άδειες χρήσης Creative Commons.
- Οι εικόνες προέρχονται από το βιβλίο «Κατανεμημένα Συστήματα με Java», Ι. Κάβουρας, Ι. Μήλης, Γ. Ξυλωμένος, Α. Ρουκουνάκη, 3^η έκδοση, 2011, Εκδόσεις Κλειδάριθμος.



Σκοποί ενότητας

- Κατανόηση του τρόπου υλοποίησης πολυεκπομπής μέσω DHT, με παράδειγμα το σύστημα Scribe.
- Εξοικείωση με τον τρόπο χρήσης των DHT για αποθήκευση ολόκληρων αρχείων, με παράδειγμα το σύστημα PAST.
- Εξοικείωση με τον τρόπο χρήσης των DHT για αποθήκευση ομάδων αρχείων, με παράδειγμα το σύστημα CFS.

Περιεχόμενα ενότητας

- Το σύστημα Scribe
- Το σύστημα PAST
- Το σύστημα CFS

**ΟΙΚΟΝΟΜΙΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ**



**ATHENS UNIVERSITY
OF ECONOMICS
AND BUSINESS**

Το σύστημα Scribe

Μάθημα: Κινητά και Διάχυτα Συστήματα, **Ενότητα # 6:** Εφαρμογές DHT

Διδάσκων: Γιώργος Ξυλωμένος, **Τμήμα:** Πληροφορικής



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



Chord και Pastry

- Ομοιότητες Chord και Pastry
 - Απεικονίζουν κόμβους / κλειδιά σε δακτύλιο
- Διαφορές Chord και Pastry
 - Σε ποιον κόμβο ανήκει ένα κλειδί;
 - Στο Chord ανήκει στον (αριθμητικά) επόμενο
 - Στο Pastry ανήκει στον (αριθμητικά) πλησιέστερο
 - Τι διεπαφή παρέχεται;
 - Στο Chord εντοπίζεται ο επόμενος
 - Στο Pastry στέλνεται μήνυμα στον πλησιέστερο

Εισαγωγή στο Scribe (1 από 4)

- Πολυεκπομπή στο Scribe
 - Αξιοποιεί δρομολόγηση με βάση το κλειδί
 - Σχεδιάστηκε για το Pastry
 - Διεπαφή δρομολόγησης μηνυμάτων
 - Μπορεί να χρησιμοποιηθεί με το Chord
 - Πρέπει όμως να τροποποιηθεί η διεπαφή

Εισαγωγή στο Scribe (2 από 4)

- Ομάδες πολυεκπομπής
 - Προσδιορίζονται από ένα κλειδί
 - Όλες οι λειτουργίες γίνονται με βάση το κλειδί
 - Δημιουργία ομάδας
 - Προσχώρηση και αποχώρηση κόμβων
 - Αποστολή δεδομένων στην ομάδα
 - Τα δεδομένα λαμβάνονται από όλα τα μέλη

Εισαγωγή στο Scribe (3 από 4)

- Κλειδιά ομάδων
 - Παράγονται με κατακερματισμό
 - MD5, SHA-1, κ.λπ.
 - Ίδιο πεδίο με τα κλειδιά των κόμβων
 - Αυθαίρετα ονόματα ομάδων
 - Ο κατακερματισμός τα κάνει ομοιόμορφα
 - Συνυπάρχουν διάφοροι τύποι ονομάτων
 - Παράδειγμα: ιεραρχικά, επίπεδα

Εισαγωγή στο Scribe (4 από 4)

- Scribe: βιβλιοθήκη πάνω από το Pastry
 - Τέσσερις βασικές κλήσεις
 - `create(key)` : δημιουργία
 - `join(key, handler)` : προσχώρηση
 - Τα μηνύματα προωθούνται στη `handler`
 - `leave(key)` : αποχώρηση
 - `multicast(key, msg)` : αποστολή
 - Κάθε κλήση παράγει μήνυμα με το ίδιο όνομα

Δημιουργία δένδρου (1 από 4)

- Σημείο ραντεβού (rendezvous) της ομάδας
 - Κόμβος με αριθμητικά πλησιέστερο κλειδί
 - Τα μηνύματα δρομολογούνται εκεί
 - `create`, `join`, `leave`, `multicast`
 - Λειτουργεί ως ρίζα δένδρου πολυεκπομπής
 - Προωθεί τα μηνύματα στους παραλήπτες
 - Μαθαίνει ότι είναι σημείο ραντεβού στο `create`

Δημιουργία δένδρου (2 από 4)

- Επεξεργασία μηνυμάτων
 - Τα μηνύματα στον παραλήπτη
 - Το μήνυμα `multicast` έχει πολλούς παραλήπτες
 - Το σημείο ραντεβού και τα μέλη της ομάδας
 - Τα `join/leave` και στους ενδιάμεσους
 - Χρησιμοποιούνται για την κατασκευή του δένδρου
 - Το `create` μόνο στο σημείο ραντεβού
 - Δημιουργεί τη ρίζα του δένδρου

Δημιουργία δένδρου (3 από 4)

- Κατάσταση κόμβων δένδρου πολυεκπομπής
 - Πίνακας προώθησης ανά ομάδα
 - Περιέχει τα παιδιά του κόμβου στο δένδρο
 - Αρχικά είναι κενός
 - Προσθήκη παιδιών: μηνύματα `join`
 - Τα μηνύματα `join` δημιουργούν το δένδρο
 - Διαγραφή παιδιών: μηνύματα `leave`
 - Τα μηνύματα `leave` καταστρέφουν το δένδρο

Δημιουργία δένδρου (4 από 4)

- Δημιουργία ομάδας
 - Μήνυμα `create` στο σημείο ραντεβού
 - Το επεξεργάζεται μόνο το σημείο ραντεβού
 - Δρομολόγηση μέσω του DHT
 - Το σημείο δημιουργεί πίνακα προώθησης
 - Αρχικά ο πίνακας είναι κενός
 - Θα προστεθούν παιδιά σταδιακά
 - Έλεγχος δικαιωμάτων δημιουργίας (προαιρετικά)

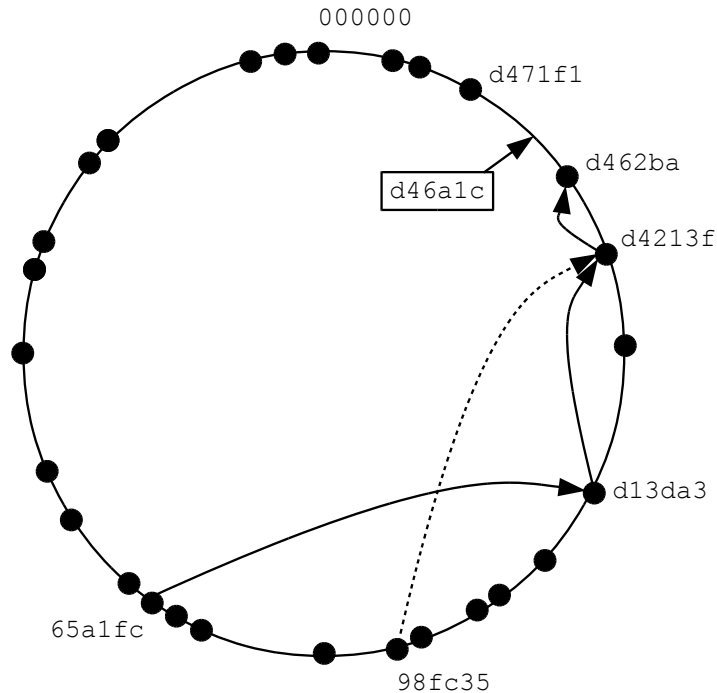
Προσχώρηση (1 από 3)

- Προσχώρηση σε ομάδα
 - Αποστολή `join` στο σημείο ραντεβού
 - Δρομολόγηση μέσω του DHT
 - Κάθε ενδιάμεσος κόμβος εξετάζει το μήνυμα
 - Αν έχει πίνακα προώθησης για την ομάδα
 - Προσθέτει τον αποστολέα στον πίνακα
 - Δεν προωθεί το `join` παραπέρα
 - Ο αποστολέας έχει συνδεθεί στο δένδρο

Προσχώρηση (2 από 3)

- Προσχώρηση σε ομάδα
 - Αν δεν έχει πίνακα προώθησης για την ομάδα
 - Δημιουργεί τον πίνακα προώθησης
 - Προσθέτει τον αποστολέα στον πίνακα
 - Αλλάζει τον αποστολέα του `join` στον εαυτό του
 - Προωθεί το `join` παραπέρα
 - Σταδιακή δημιουργία δένδρου πολυεκπομπής
 - Από τα φύλλα προς το σημείο ραντεβού

Προσχώρηση (3 από 3)



- Παράδειγμα προσχώρησης
 - Ομάδα `d46a1c`, προσχωρούν `65a1fc` και `98fc35`
 - Το δεύτερο `join` δεν φτάνει στο σημείο ραντεβού
 - Σταματάει στον πρώτο κόμβο που είναι ήδη στο δένδρο

Αποχώρηση

- Αποχώρηση από ομάδα
 - Αποστολή `leave` στο σημείο ραντεβού
 - Κάθε ενδιάμεσος κόμβος εξετάζει το μήνυμα
 - Αφαίρεση αποστολέα από πίνακα προώθησης
 - Αν ο πίνακα προώθησης δεν μένει κενός
 - Δεν προωθεί το `leave` παραπέρα
 - Αν ο πίνακα προώθησης μένει κενός
 - Αλλάζει τον αποστολέα του `leave` στον εαυτό του
 - Προωθεί το `leave` παραπέρα

Αποστολή μηνυμάτων

- Αποστολή μηνυμάτων στην ομάδα
 - Αποστολή `multicast` στο σημείο ραντεβού
 - Το σημείο ραντεβού προωθεί το μήνυμα
 - Χρήση πάντα του DHT
 - Αναδρομική προώθηση μέχρι τα φύλλα
 - Σε κάθε μέλος καλείται η μέθοδος `handler`
 - Ουσιαστικά είναι η μέθοδος λήψης μηνυμάτων
 - Μπορεί να περάσουμε πολλά μέλη στη διαδρομή

Απόδοση (1 από 2)

- Πόσο αποδοτική είναι η πολυεκπομπή;
 - Το δένδρο είναι ένωση μονοπατιών
 - Τα μονοπάτια είναι αντίστροφα των επιθυμητών
 - Από τα μέλη της ομάδας προς το σημείο ραντεβού
 - Μπορεί να μην είναι συμμετρικά
 - Τα μονοπάτια προς τα μέλη δεν είναι βέλτιστα
 - Ως προς το υποκείμενο δίκτυο
 - Το Pastry προσεγγίζει τα τοπολογικά βέλτιστα

Απόδοση (2 από 2)

- Πόσο αποδοτική είναι η πολυεκπομπή;
 - Το σημείο ραντεβού δεν είναι βέλτιστο
 - Τα μηνύματα κινούνται πρώτα προς σημείο ραντεβού
 - Μπορεί να απομακρύνονται από τα μέλη της ομάδας
 - Το μονοπάτι προς το ραντεβού δεν είναι βέλτιστο
 - Ο αποστολέας μπορεί να αποθηκεύσει τη διεύθυνση
 - Στη συνέχεια αποστολή μηνυμάτων μέσω IP

Αποτυχίες (1 από 2)

- Αποτυχία / αποχώρηση εσωτερικών κόμβων
 - Κάθε κόμβος στέλνει περιοδικά μηνύματα
 - Σε όλα τα παιδιά του
 - Αν ένα παιδί δεν λάβει μήνυμα για αρκετή ώρα
 - Υποθέτει ότι ο πατέρας του έχει αποτύχει
 - Στέλνει νέο `join` προς το σημείο ραντεβού
 - Το παιδί επανασυνδέεται με το δένδρο

Αποτυχίες (2 από 2)

- Αποτυχία ή αποχώρηση σημείου ραντεβού
 - Το σημείο ραντεβού αναπαράγει την κατάστασή του
 - Περιοδική αποστολή κατάστασης στο σύνολο φύλλων
 - Αν αποτύχει, κάποιο φύλλο θα γίνει σημείο ραντεβού
 - Τα παιδιά παρατηρούν την αποτυχία
 - Δεν λαμβάνουν μήνυμα από τον πατέρα
 - Στέλνουν νέο μήνυμα `join` προς το σημείο ραντεβού
 - Τα μηνύματα `join` καταλήγουν στο νέο σημείο ραντεβού
 - Παρόμοια αντιμετώπιση από τους αποστολείς

**ΟΙΚΟΝΟΜΙΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ**



**ATHENS UNIVERSITY
OF ECONOMICS
AND BUSINESS**

Το σύστημα PAST

Μάθημα: Κινητά και Διάχυτα Συστήματα, **Ενότητα # 6:** Εφαρμογές DHT

Διδάσκων: Γιώργος Ξυλωμένος, **Τμήμα:** Πληροφορικής



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



Εισαγωγή στο PAST (1 από 5)

- PAST: αποθήκευση αρχείων μέσω Pastry
 - Αντιστοίχιση κάθε αρχείου σε k κόμβους
 - Αριθμητικά πλησιέστεροι στο κλειδί του αρχείου
 - Βαθμός αναπαραγωγής k
 - Αύξηση αξιοπιστίας και αποδοτικότητας
 - Δεν υποστηρίζει πλήρες σύστημα αρχείων
 - Απλά αποθηκεύει τα αρχεία αυτόνομα

Εισαγωγή στο PAST (2 από 5)

- Διεπαφή προσανατολισμένη σε αρχεία
 - Βασίζεται σε σύστημα δημόσιας κρυπτογραφίας
 - Θεωρούμε ότι γνωρίζουμε τα δημόσια κλειδιά
 - `key=insert(name, k, file)`
 - Αποθήκευση `file` με όνομα `name` σε `k` κόμβους
 - Το `key` παράγεται με κατακερματισμό των στοιχείων
 - Όνομα, δημόσιο κλειδί και τυχαία επιλεγμένη τιμή

Εισαγωγή στο PAST (3 από 5)

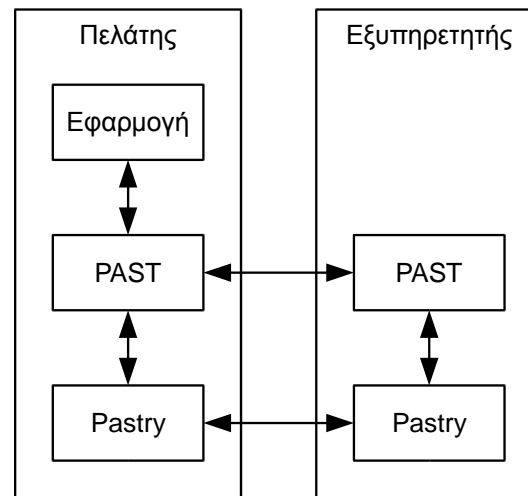
- Διεπαφή προσανατολισμένη σε αρχεία
 - `file=lookup(key)`
 - Ανάκτηση ενός αντιγράφου του αρχείου με κλειδί `key`
 - Συνήθως από τον τοπολογικά πλησιέστερο κόμβο
 - `reclaim(key)`
 - Απελευθέρωση του χώρου του αρχείου με κλειδί `key`
 - Δεν εγγυάται ότι θα διαγραφούν όλα τα αντίγραφα

Εισαγωγή στο PAST (4 από 5)

- Ιδιαιτερότητες της διεπαφής του PAST
 - Δεν επιτρέπει την τροποποίηση των αρχείων
 - Αν εισάγουμε ξανά το αρχείο θα αλλάξει κλειδί
 - Θα χρησιμοποιηθεί διαφορετική τυχαία τιμή
 - Τα αρχεία είναι αμετάβλητα (immutable)
 - Μόνο ο ιδιοκτήτης γνωρίζει το κλειδί του αρχείου
 - Αυτός έχει την ευθύνη και για τη διανομή του
 - Δεν γίνεται έλεγχος πρόσβασης στα lookups
 - Έλεγχος μόνο στη διαγραφή των αρχείων

Εισαγωγή στο PAST (5 από 5)

- Μοντέλο ανεβάσματος-κατεβάσματος
 - Παρόμοιο με το AFS αλλά χωρίς σύστημα αρχείων
 - Η εφαρμογή βλέπει τη διεπαφή του PAST
 - Η εφαρμογή δεν τρέχει σε όλους τους κόμβους



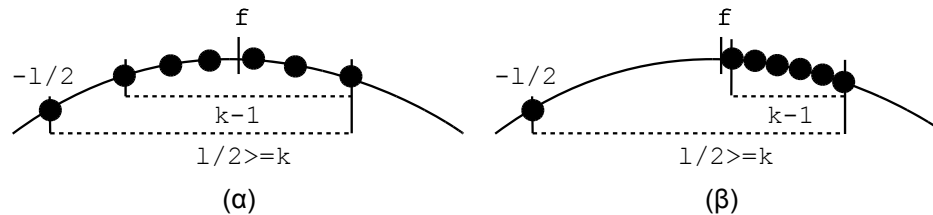
Εισαγωγή αρχείων (1 από 3)

- Εισαγωγή αρχείων στο PAST
 - Υπολογισμός κλειδιού με κατακερματισμό
 - Όνομα αρχείου, δημόσιο κλειδί, τυχαία τιμή
 - Δημιουργία πιστοποιητικού αρχείου
 - Κλειδί, hash περιεχομένου, τιμή k , τυχαία τιμή, ώρα
 - Δεν μεταβάλλεται αφού το αρχείο δεν αλλάζει
 - Υπογράφεται ψηφιακά από τον ιδιοκτήτη
 - Αποθηκεύεται μαζί με το αρχείο

Εισαγωγή αρχείων (2 από 3)

- Εισαγωγή αρχείων στο PAST
 - Αρχείο και πιστοποιητικό δρομολογούνται
 - Όταν φτάσουν σε έναν από τους k κόμβους
 - Ελέγχεται η εγκυρότητα αρχείου και πιστοποιητικού
 - Το αρχείο αποθηκεύεται
 - Το αρχείο στέλνεται στους $k-1$
 - Κάθε παραλήπτης επιστρέφει απόδειξη αποθήκευσης
 - Ο πελάτης θα λάβει k έγκυρες αποδείξεις

Εισαγωγή αρχείων (3 από 3)



- Εντοπισμός των k πλησιέστερων κόμβων
 - Τα μηνύματα δρομολογούνται προς το κλειδί f
 - Κάθε κόμβος διατηρεί ένα σύνολο 1 φύλλων
 - Οι $1/2$ μικρότεροι και $1/2$ μεγαλύτεροι κόμβοι
 - Απαιτούμε $1/2 \geq k$
 - Έστω ότι το μήνυμα φτάνει στον ακραίο κόμβο
 - Οι $k-1$ άλλοι κόμβοι βρίσκονται στο σύνολο φύλλων
 - Άρα ο πρώτος γνωρίζει πάντα τους άλλους $k-1$

Διαγραφή αρχείων

- Διαγραφή αρχείων στο PAST
 - Δημιουργία πιστοποιητικού ανάκτησης
 - Υπογράφεται ψηφιακά από τον ιδιοκτήτη
 - Η αίτηση και το πιστοποιητικό δρομολογούνται
 - Όταν φτάσουν σε έναν από τους k κόμβους
 - Ελέγχεται η εγκυρότητα του πιστοποιητικού
 - Το αρχείο διαγράφεται
 - Η αίτηση δρομολογείται στους $k-1$
 - Κάθε παραλήπτης επιστρέφει μια απόδειξη διαγραφής
 - Ο πελάτης θα λάβει k έγκυρες αποδείξεις (ίσως!)

Ανάκτηση αρχείων

- Ανάκτηση αρχείων στο PAST
 - Το μήνυμα αποστέλλεται μέσω του Pastry
 - Ο πρώτος από τους k κόμβους απαντά
 - Επιστρέφεται και το πιστοποιητικό αρχείου
 - Ο πελάτης επιβεβαιώνει την εγκυρότητα του αρχείου
- Χρήση πιστοποιητικών
 - Το PAST δημιουργεί πιστοποιητικά
 - Κατά την εισαγωγή και διαγραφή αρχείων
 - Στις άλλες λειτουργίες απλά τα επιβεβαιώνει

Αποθήκευση σε πολλά σημεία

- Κάθε αρχείο αποθηκεύεται σε k κόμβους
 - Οι κόμβοι είναι διάσπαρτοι στο σύστημα
 - Μικρή πιθανότητα ταυτόχρονης αποτυχίας
 - Κάποιος μπορεί να είναι κοντά στον πελάτη
- Προσωρινή αποθήκευση αρχείων
 - Κάθε κόμβος αξιοποιεί τον ελεύθερο χώρο του
 - Αποθηκεύει προσωρινά διερχόμενα αρχεία
 - Διαχείριση χώρου με κάποια πολιτική
 - Προτεραιότητα έχουν τα αντίγραφα αρχείων

Κατανομή φόρτου

- Βασίζεται στην τυχαία κατανομή αρχείων
 - Αυτό δεν οδηγεί σε ομοιόμορφη κατανομή
- Ανομοιόμορφες δυνατότητες κόμβων
 - Ορισμένοι κόμβοι έχουν πολύ διαθέσιμο χώρο
 - Οι κόμβοι διασπώνται σε εικονικούς κόμβους
 - Μοιράζονται το φυσικό χώρο αποθήκευσης
- Ανομοιόμορφη κατανομή αρχείων
 - Αρχεία με πάρα πολύ μεγάλο μέγεθος
 - Κάποιοι κόμβοι μπορεί να φορτωθούν περισσότερο

Εκτροπή αντιγράφων (1 από 4)

- Εκτροπή αντιγράφων
 - Έστω ότι ο A λαμβάνει μία αίτηση εισαγωγής
 - Αν δεν μπορεί να την ικανοποιήσει;
 - Επιλέγει ένα από τα φύλλα του, τον B
 - Δεν πρέπει να είναι στους k πλησιέστερους
 - Δεν πρέπει να έχει δεχτεί ήδη άλλο αντίγραφο
 - Στέλνει το αρχείο στον B
 - Αποθηκεύει έναν δείκτη προς το αντίγραφο
 - Εκδίδει την απόδειξη αποθήκευσης

Εκτροπή αντιγράφων (2 από 4)

- Εκτροπή αντιγράφων
 - Περιοδικά οι δείκτες στέλνονται και αλλού
 - Για κάθε δείκτη εντοπίζεται ο $k+1$ πλησιέστερος
 - Έστω ότι είναι ο C
 - Ο δείκτης εκτροπής στέλνεται και στον C
 - Αν αποτύχει ο A , ο C παίρνει αυτόματα τη θέση του
 - Αν αποτύχει ο C , ο A εντοπίζει κάποιον άλλο κόμβο

Εκτροπή αντιγράφων (3 από 4)

- Κριτήρια εκτροπής αντιγράφων
 - Κάθε αίτηση εισαγωγής εξετάζεται χωριστά
 - Εκτρέπονται αρχεία που θέλουν πολύ χώρο
 - Πάνω από κάποιο ποσοστό του διαθέσιμου χώρου
 - Δεν γίνεται από υποφορτωμένους κόμβους
 - Εκτρέπονται κυρίως μεγάλα αρχεία

Εκτροπή αντιγράφων (4 από 4)

- Αν δεν δεχτεί κανένα φύλλο το αρχείο;
 - Μπορεί όλα τα φύλλα να είναι φορτωμένα
 - Χρήση παρόμοιας πολιτικής με αυτή της εκτροπής
 - Αναγκαστικά γίνεται εκτροπή αρχείου
 - Επιστροφή σφάλματος στον πελάτη
 - Ο πελάτης δημιουργεί νέο κλειδί για το αρχείο
 - Απλά αλλάζει την τυχαία τιμή
 - Το αρχείο αποθηκεύεται με το νέο κλειδί
 - Το παλιό κλειδί διαγράφεται

Διαχείριση αλλαγών (1 από 3)

- Αποτυχία, αποχώρηση, προσχώρηση κόμβων
 - Αλλάζει το σύνολο φύλλων ορισμένων κόμβων
 - Αν αλλάξουν οι k πλησιέστεροι ενός αρχείου;
 - Δημιουργία νέων αντιγράφων του αρχείου
 - Πρέπει να έχουμε k στους σωστούς κόμβους
 - Περιλαμβάνονται οι κόμβοι με εκτροπή αντιγράφων
 - Και αυτοί ανήκουν στο σύνολο φύλλων ενός εκ των k

Διαχείριση αλλαγών (2 από 3)

- Αποτυχία ή αποχώρηση
 - Ένα κόμβος μπορεί να χρειάζεται αντίγραφα
 - Οι $k-1$ θα βρίσκονται στο σύνολο φύλλων του
 - Αρκεί να ρωτήσει τα φύλλα του ποια αρχεία έχουν
 - Μετά αποφασίζει ποια αρχεία του αντιστοιχούν
 - Τέλος ζητάει τα αρχεία από αυτούς τους κόμβους

Διαχείριση αλλαγών (3 από 3)

- Προσχώρηση κόμβων
 - Το σύστημα περιέχει ήδη k αντίγραφα
 - Αρκεί ο νέος κόμβος να προσθέσει δείκτες
 - Αρχικά εντοπίζει τον κόμβο που έφυγε από τους k
 - Στη συνέχεια αποθηκεύει δείκτες προς αυτόν
 - Σταδιακά τα αντίγραφα έρχονται στον κόμβο
 - Τα αντίγραφα μπορεί να απομακρυνθούν από τους k
 - Αν έχουμε αποτυχίες θα μείνουν λιγότερα από k

**ΟΙΚΟΝΟΜΙΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ**



**ATHENS UNIVERSITY
OF ECONOMICS
AND BUSINESS**

Το σύστημα CFS

Μάθημα: Κινητά και Διάχυτα Συστήματα, **Ενότητα # 6:** Εφαρμογές DHT

Διδάσκων: Γιώργος Ξυλωμένος, **Τμήμα:** Πληροφορικής



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

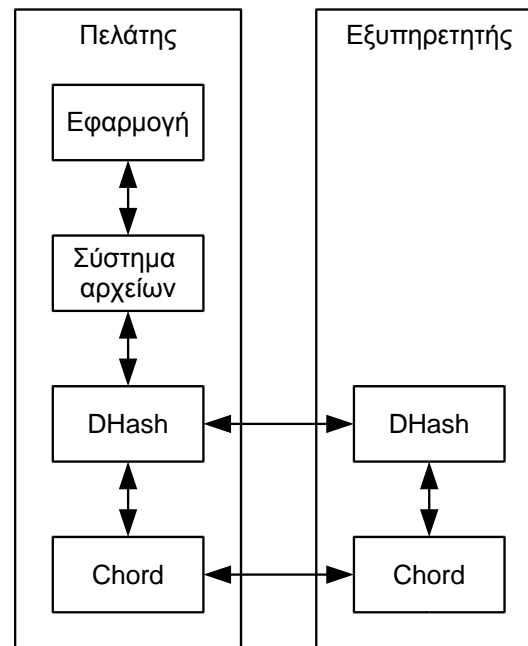
Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Εισαγωγή στο CFS (1 από 3)

- CFS: Αποθήκευση αρχείων μέσω Chord
 - Αντιστοίχιση ομάδας δεδομένων σε k κόμβους
 - Βαθμός αναπαραγωγής k
 - Υποστηρίζει πλήρες σύστημα αρχείων



Εισαγωγή στο CFS (2 από 3)

- Δομή του CFS
 - Το Chord απεικονίζει τα κλειδιά ομάδων σε κόμβους
 - Το DHash διαχειρίζεται τις ομάδες των δεδομένων
 - Το σύστημα αρχείων συνθέτει αρχεία και ευρετήρια
- Ομοιότητες με το PAST
 - Εικονικοί κόμβοι για εξισορρόπηση φόρτου
- Διαφορές με το PAST
 - Λειτουργία σε επίπεδο ομάδων και όχι αρχείων
 - Πιο ομοιόμορφη κατανομή δεδομένων
 - Απλούστερη διαχείριση αντιγράφων

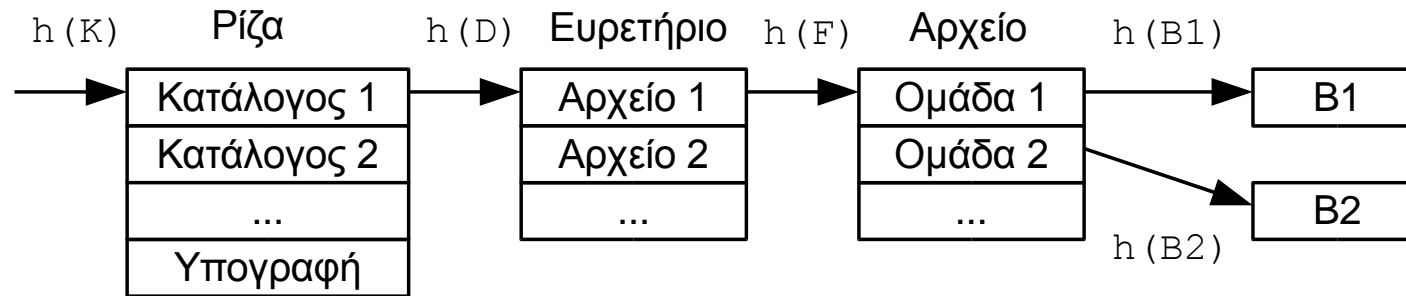
Εισαγωγή στο CFS (3 από 3)

- Διεπαφή αρχείων του CFS
 - Παρόμοια με αυτή του UNIX
 - Υλοποιεί μοντέλο απομακρυσμένης πρόσβασης
 - Μόνο ο ιδιοκτήτης μπορεί να αλλάξει τα αρχεία
 - Για τους πελάτες τα αρχεία είναι ανάγνωσης μόνο
 - Οι ομάδες αποθηκεύονται για περιορισμένο χρόνο
 - Ο ιδιοκτήτης πρέπει να τις ανανεώνει
 - Αλλιώς οι ομάδες διαγράφονται
 - Δεν υπάρχει ρητή κλήση διαγραφής αρχείων
 - Υποθέτουμε σύστημα δημόσιας κρυπτογραφίας

Σύστημα αρχείων (1 από 3)

- Ξεκινάμε με τις ομάδες δεδομένων $B1$ και $B2$
 - Αποθηκεύονται με κλειδιά $h(B1)$ και $h(B2)$
- Το αρχείο F περιέχει τα κλειδιά των ομάδων
 - Αντίστοιχο με το i -node στο UNIX
 - Αποθηκεύεται με κλειδί $h(F)$
- Το ευρετήριο D περιέχει τα κλειδιά των αρχείων
 - Αποθηκεύεται με κλειδί $h(D)$
- Το ευρετήριο ρίζας υπογράφεται από τον ιδιοκτήτη
 - Αποθηκεύεται με κλειδί $h(K)$ (K είναι το δημόσιο κλειδί)
 - Ένα σύστημα αρχείων ανά δημόσιο κλειδί

Σύστημα αρχείων (2 από 3)



- Έλεγχος δεδομένων
 - Κάθε ομάδα ταιριάζει με το κλειδί της
 - Η ρίζα ελέγχεται με βάση την υπογραφή
 - Πρέπει να ταιριάζει με το κλειδί του ευρετηρίου
 - Οι ομάδες ελέγχονται με βάση το περιεχόμενο
 - Πρέπει να ταιριάζουν με τα κλειδιά τους

Σύστημα αρχείων (3 από 3)

- Τροποποίηση αρχείων
 - Κάθε αλλαγή σε ομάδα αλλάζει το κλειδί της
 - Προέρχεται από κατακερματισμό της ομάδας
 - Τελικά αυτό απαιτεί αλλαγή της ρίζας
 - Το νέο κλειδί αλλάζει την ομάδα που καταχωρείται
 - Μόνο ο ιδιοκτήτης αλλάζει τα δεδομένα
 - Καταχωρεί νέα ρίζα και την υπογράφει ψηφιακά
 - Πρέπει να χρησιμοποιήσει το ίδιο δημόσιο κλειδί

Διεπαφή του CFS (1 από 2)

- Η διαχείριση ομάδων γίνεται από το DHash
 - Το Chord χρησιμοποιείται για τις αναζητήσεις
 - Το DHash αξιοποιείται από το σύστημα αρχείων
- Διεπαφή του DHash
 - `put_h(block)`
 - Υπολογίζει το κλειδί κατακερματίζοντας το `block`
 - Στέλνει την ομάδα για αποθήκευση σε `k` κόμβους

Διεπαφή του CFS (2 από 2)

- Διεπαφή του DHash

- `put_s(block, pubkey)`

- Επιβεβαιώνει ότι η `block` είναι έγκυρη

- Πρέπει να είναι ψηφιακά υπογεγραμμένη με `pubkey`

- Χρησιμοποιείται μόνο για το ριζικό ευρετήριο

- Υπολογίζει κλειδί με κατακερματισμό του `pubkey`

- Στέλνει την ομάδα για αποθήκευση σε `k` κόμβους

- `block=get(key)`

- Ανακτά την ομάδα που προσδιορίζεται από το `key`

Αποθήκευση ομάδων (1 από 2)

- Αποθήκευση ομάδων σε κόμβους
 - Το DHash εντοπίζει το διάδοχο της ομάδας
 - Ο κόμβος αποθηκεύει την ομάδα σε k διαδόχους
 - Οι διάδοχοι βρίσκονται στη λίστα διαδόχων
 - Η λίστα διαδόχων είναι μεγαλύτερη του k
 - Αν αποτύχει αναλαμβάνει ο νέος διάδοχος
 - Έχει ήδη αποθηκευμένη την ομάδα

Αποθήκευση ομάδων (2 από 2)

- Τυχαία κατανομή των αντιγράφων
 - Οι k κόμβοι είναι διάσπαρτοι στο δίκτυο
 - Πολύ απίθανο να αποτύχουν ταυτόχρονα
 - Στις ανακτήσεις ο διάδοχος επιστρέφει τους k
 - Ο πελάτης επιλέγει τον τοπολογικά πλησιέστερο
 - Αρκετά πιθανό κάποιος από τους k να είναι κοντά

Προσωρινή αποθήκευση

- Προσωρινή αποθήκευση ομάδων
 - Όταν γίνεται αναζήτηση οι ενδιάμεσοι κόμβοι ερωτώνται
 - Αν περιέχουν την ομάδα, την επιστρέφουν άμεσα
 - Αλλιώς θα τη λάβουν από τον πελάτη όταν την ανακτήσει
 - Δεν απαιτείται έλεγχος εγκυρότητας των ομάδων
 - Το κλειδί τους πιστοποιεί αν είναι έγκυρες
 - Αν αλλάξουν, αλλάζει και το κλειδί τους
 - Εξαιρείται το ευρετήριο ρίζας
 - Το κλειδί του δεν αλλάζει όταν τροποποιείται
 - Πρέπει να εξετάζεται πόσο πρόσφατο είναι

Ανανέωση αρχείων

- Περιοδική ανανέωση αρχείων
 - Στο CFS οι ομάδες δεν διαγράφονται ρητά
 - Απλά διαγράφονται όταν παλιώσουν πολύ
 - Το σύστημα αρχείων διασχίζεται περιοδικά
 - Ξεκινάμε από τη ρίζα και φτάνουμε στα φύλλα
 - Κάθε ομάδα που εντοπίζεται αποθηκεύεται ξανά
 - Οι ομάδες που έχουν διαγραφεί δεν ανανεώνονται

**ΟΙΚΟΝΟΜΙΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ**



**ATHENS UNIVERSITY
OF ECONOMICS
AND BUSINESS**

Τέλος Ενότητας #6

Μάθημα: Κινητά και Διάχυτα Συστήματα, **Ενότητα # 6:** Εφαρμογές DHT

Διδάσκων: Γιώργος Ξυλωμένος, **Τμήμα:** Πληροφορικής



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

