

Κινητός και Διάχυτος Υπολογισμός (Mobile & Pervasive Computing)

Δημήτριος Κατσαρός

Χειμώνας 2015

Διάλεξη 4η

Περιεχόμενα

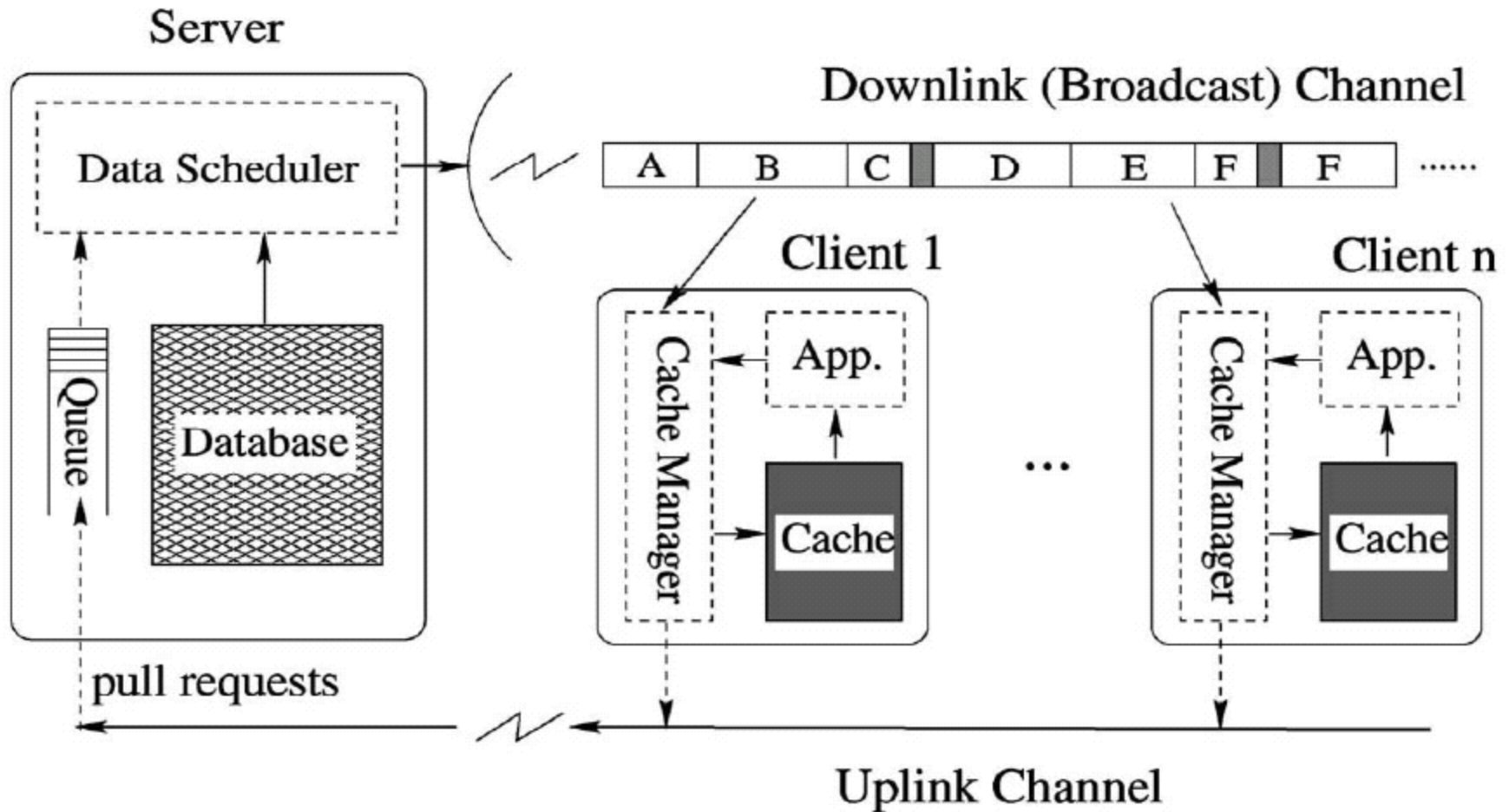
- **Caching**
- Prefetching

Caching στους κινητούς πελάτες

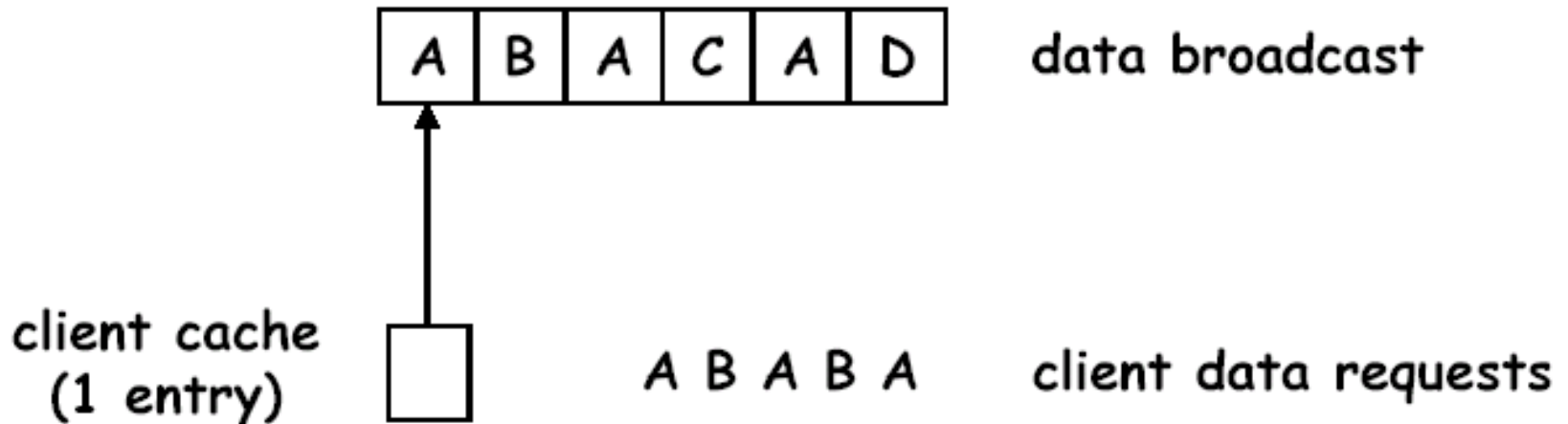
- Προγράμματα εκπομπής
 - Βασίζονται στη μέση πιθανότητα προσπέλασης: μέσος όρος πάνω σε ΌΛΟΥΣ τους πελάτες
 - Όχι αναγκαστικά βέλτιστη για κάθε έναν πελάτη
- Πώς μπορεί κάθε πελάτης να υποβοηθήσει τον εαυτό του?
 - **Caching**: προσωρινή αποθήκευση των δεδομένων που λαμβάνει
 - Πολιτική caching: όταν εξαναγκάζεται να αντικαταστήσει κάποιο (επειδή η cache είναι πλήρης), αντικαθιστά εκείνα που είναι λιγότερο πιθανό να χρησιμεύσουν στο μέλλον

Το γενικό μοντέλο caching

□ Data Item ■ Invalidation Report

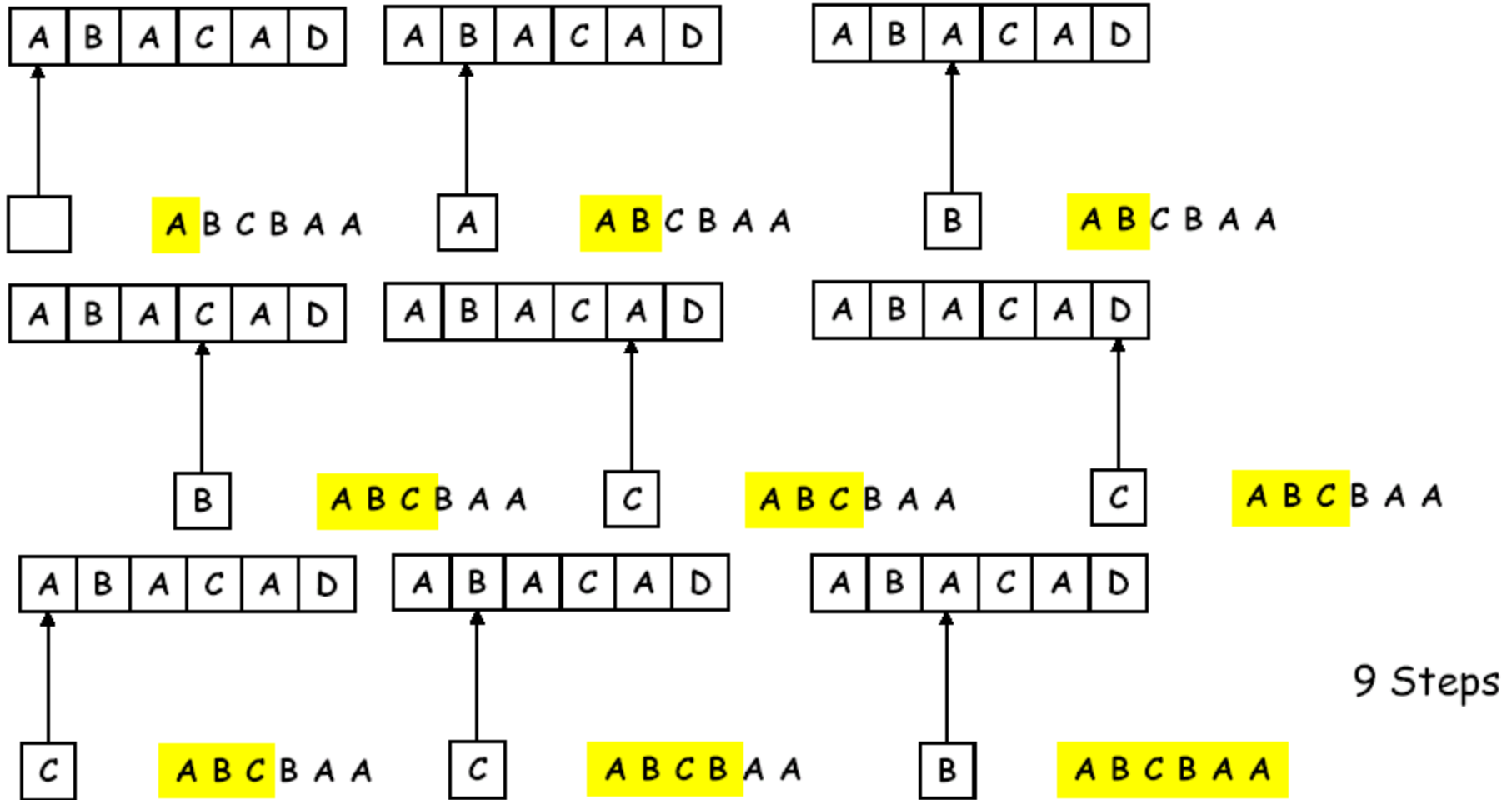


Caching στους κινητούς πελάτες



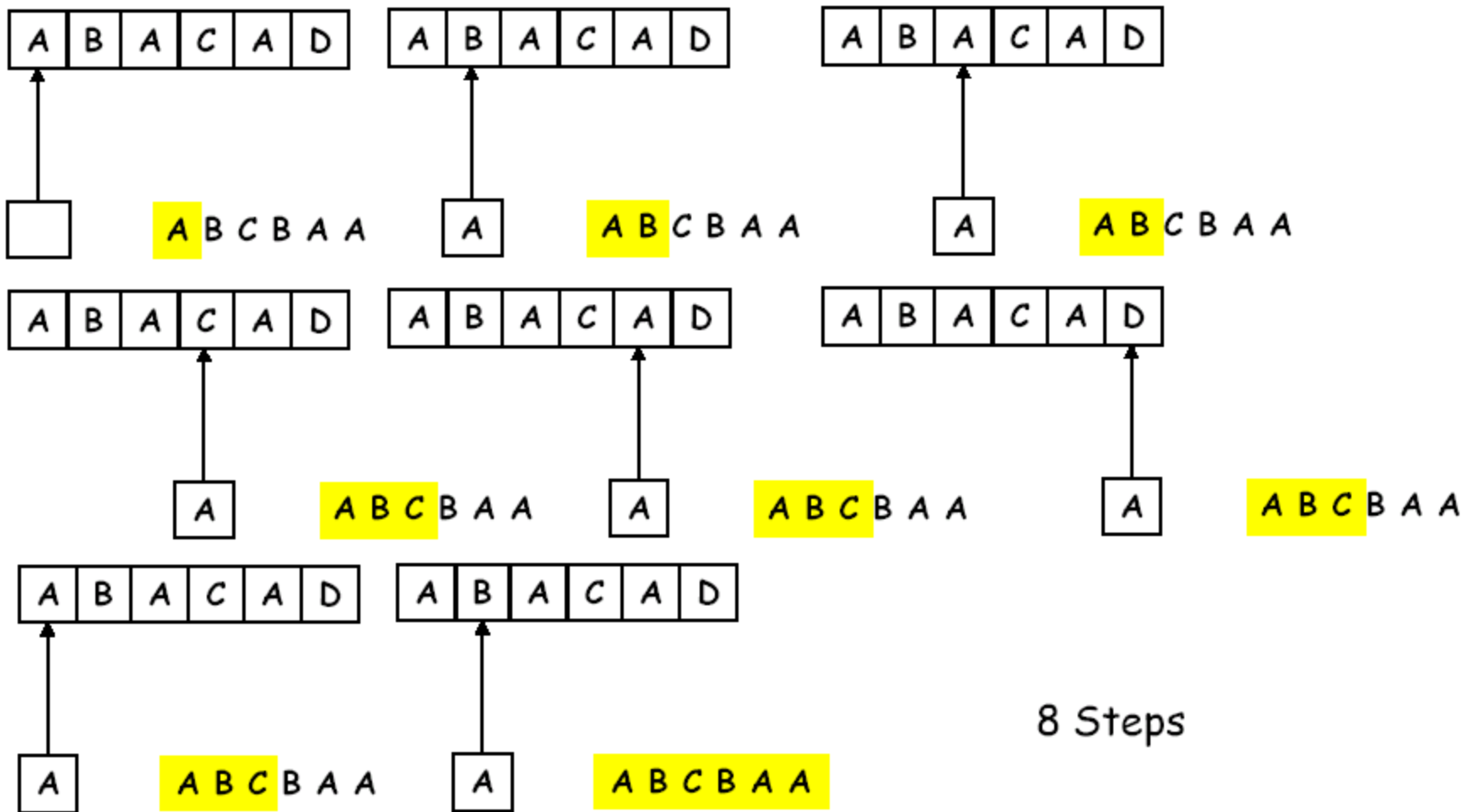
- Ποια **πολιτική αντικατάστασης** (replacement strategy ή policy ή algorithm) θα πρέπει να χρησιμοποιούν οι πελάτες;

Least Recently Used (LRU)



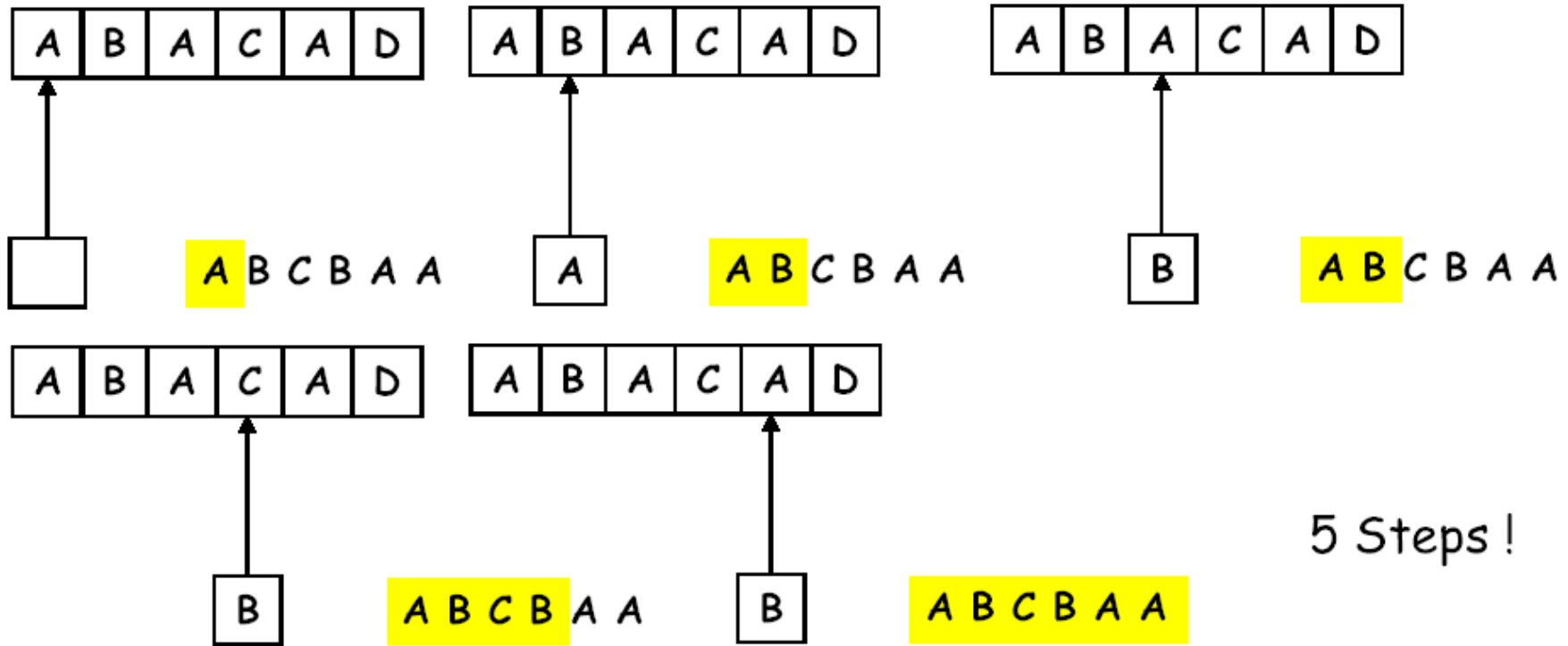
LRU: Αντικατάσταση του αντικειμένου που έχει χρησιμοποιηθεί παλιότερα στο παρελθόν

Most Probable Accessed (MPA)



MPA: Αντικατάσταση του αντικειμένου που θα χρησιμοποιηθεί λιγότερο συχνά στο μέλλον

Probability Inv. Broadc. Frequency



5 Steps !

PIX: Αντικατάσταση του αντικειμένου με το μικρότερο P/X, όπου

- **P:** είναι η πιθανότητα προσπέλασης του αντικειμένου
- **X:** είναι η συχνότητα εκπομπής του αντικειμένου

$$PIX(A) = \frac{1}{2} / \frac{1}{2} = 1, PIX(B) = \frac{1}{3} / \frac{1}{6} = 2, PIX(C) = \frac{1}{6} / \frac{1}{6} = 1$$

Η πολιτική LIX (1/3)

- Η πολιτική PIX δεν είναι πρακτική
 - Απαιτεί τέλεια γνώση των πιθανοτήτων προσπέλασης
 - Απαιτεί σύγκριση των τιμών PIX όλων των αντικειμένων στην cache (σειριακή σάρωση των αντικειμένων)
- Ιδέα: προσέγγιση της πολιτικής PIX με έναν αλγόριθμο του τύπου LRU (δηλ., τον LIX) που λαμβάνει υπόψη του πιθανότητες προσπέλασης
- LRU (Least Recently Updated)
 - Τα Cached δεδομένα διατηρούνται σε μια λίστα
 - Εάν ένα δεδομένο προσπελαστεί μετακινείται στην κορυφή/αρχή της λίστας
 - Όταν συμβεί cache miss (απαιτείται να “κατεβεί” από το κανάλι και να μπει στην cache), το δεδομένο στο τέλος της λίστας εκδιώκεται από τη λίστα

Η πολιτική LIX (2/3)

- Χρήση τροποποιημένου LRU για κάθε δίσκο εκπομπής χωριστά
 - Ο LIX διατηρεί μια λίστα με αντικείμενα για κάθε δίσκο εκπομπής με συχνότητα f_j
 - Τα αντικείμενα μπαίνουν στη λίστα του αντίστοιχου δίσκου εκπομπής όπου ανήκουν
 - Όταν ένα αντικείμενο προσπελάζεται, τοποθετείται στην κορυφή/αρχή της λίστας
 - Μια εκτίμηση της πιθανότητας προσπέλασης p_i αναπροσαρμόζεται οποτεδήποτε το αντικείμενο d_i προσπελάζεται
 - Όταν ένα αντικείμενο d_i πρόκειται να εκδιωχτεί, υπολογίζεται μια LIX τιμή lix_i (που προσεγγίζει την PIX τιμή) για κάθε αντικείμενο στο τέλος κάθε λίστας
 - Το αντικείμενο με τη χαμηλότερη τιμή, τελικά εκδιώκεται
The data item with lowest value lix_i is evicted

Η πολιτική LIX (3/3)

Computing the access probability estimate p_i for d_i :

Initial value $p_i = 0$

If most recent access to d_i was at time t_i (initially $t_i = 0$) and t is the time of the current access then

$$p_i := \frac{c}{t - t_i} + (1 - c)p_i$$

c constant ($0 < c < 1$)

Approximation of PIX value
(d_i belongs to broadcast disk with frequency f_j)

$$lix_i = p_i / f_j$$

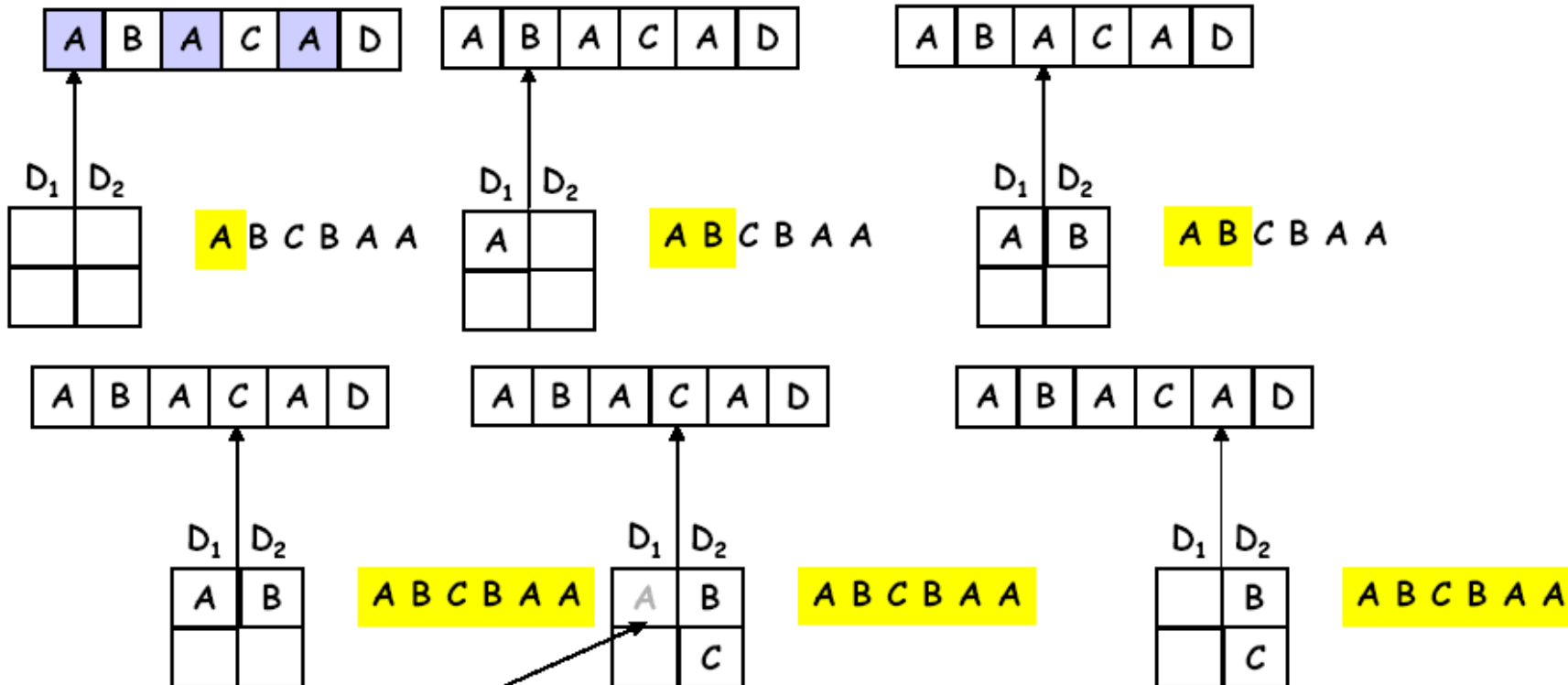
Παράδειγμα LIX

Δυο δίσκοι εκπομπής

$D_1 = \{A\}$, $D_2 = \{B, C, D\}$, συχνότητες $f_1 = 3$, $f_2 = 1$, $c = 1/2$

Η cache μπορεί να αποθηκεύσει 2 αντικείμενα.

p_i	1	2	3	4	5
A	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
B	0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$
C	0	0	0	$\frac{1}{8}$	$\frac{1}{8}$
D	0	0	0	0	0



$$\text{evict: } lix_A = \frac{1}{2}/3 = 1/6, lix_B = \frac{1}{4}/1 = 1/4$$

Κινητός και Διάχυτος Υπολογισμός (Mobile & Pervasive Computing)

Δημήτριος Κατσαρός

Χειμώνας 2015

Διάλεξη 5η

Περιεχόμενα

- **Prefetching**

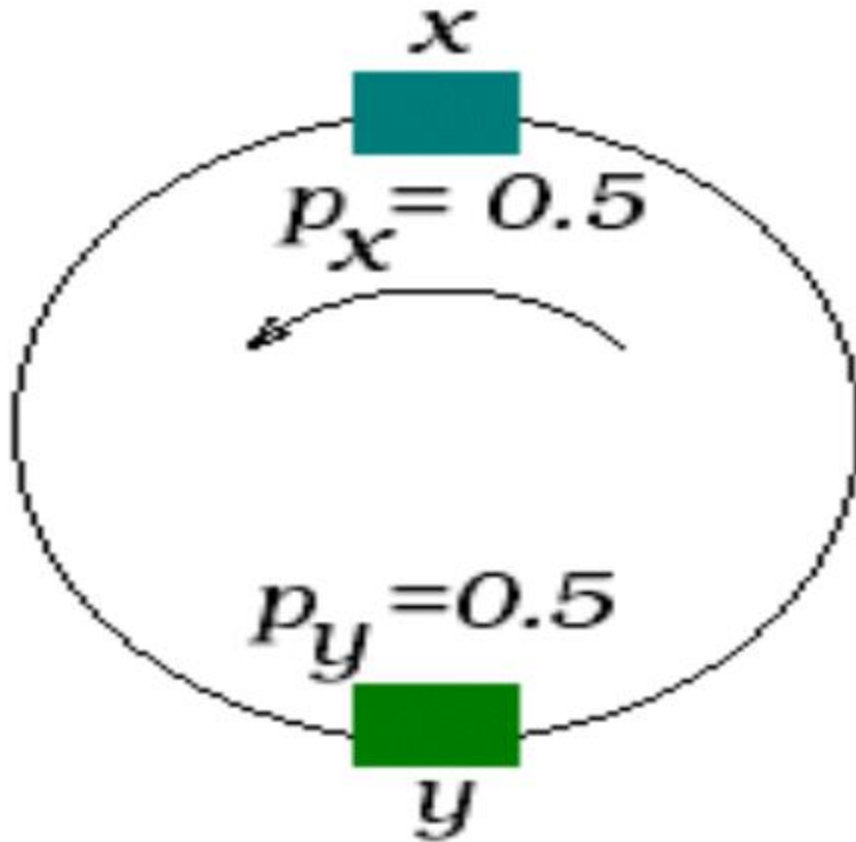
Κίνητρο για prefetching

- Οι PIX/LIX αποθηκεύουν στην cache αντικείμενα, μόνο μετά αφού ζητηθούν
- Εναλλακτικά, ο κινητός πελάτης μπορεί να “κατεβάζει” αντικείμενα από το κανάλι, αφού ούτως ή άλλως το ακούει
- Ο στόχος είναι να ελαττώσει το χρόνο απόκρισης
- Μέθοδοι prefetching:
 - **Tag Team Caching**
 - **Ευρεστικό Prefetching**

Η έννοια του tag team caching

- Tag Team Caching – Τα αντικείμενα συνεχώς αντικαθιστούν το ένα το άλλο μέσα στην cache
- Για παράδειγμα,
 - έστωσαν δυο αντικείμενα x και y τα οποία εκπέμπονται από το κανάλι
 - Ο πελάτης caches το x όταν εκπέμπεται στο κανάλι
 - Εκδιώχνει το x και caches το y , όταν εκπέμπεται το y

Παράδειγμα tag team caching (1/2)



Cache
 x

Αναμεν. Καθυστέρ. “Demand Driven”

- Υποθέτουμε ότι ένας πελάτης ενδιαφέρεται να προσπελάσει το x και το y $p_x = p_y = 0.5$ και ότι έχει μια cache με μια μόνο θέση
- Στο μοντέλο demand driven, κάνει cache το x και εάν χρειάζεται το y , περιμένει για το y και αντικαθιστά το x στην cache με το y
- Η αναμενόμενη καθυστέρηση σε cache miss είναι $\frac{1}{2}$ της περιστροφής του δίσκου
- Η αναμενόμενη καθυστέρηση πάνω σε όλες τις προσπελάσεις είναι
- $C_i * M_i * D_i$, όπου C είναι η πιθανότητα προσπέλασης, M είναι η πιθανότητα ενός cache miss και D είναι η αναμενόμενη καθυστέρηση εκπομπής για το αντικείμενο i
- Για τα αντικείμενα x και y , είναι ίση με: $0.5 * 0.5 * 0.5 + 0.5 * 0.5 * 0.5 = \mathbf{0.25}$

Αναμεν. Καθυστέρ. “Tag team caching”

- $0.5 * 0.5 * 0.25 + 0.5 * 0.5 * 0.25 = \mathbf{0.125}$, δηλαδή, το μέσο κόστος είναι το $\frac{1}{2}$ του αντίστοιχου κόστους στο σχήμα demand driven
- Γιατί? Ένα miss μπορεί να συμβεί σε οποιαδήποτε στιγμή στο μοντέλο demand driven, ενώ τα misses συμβαίνουν μόνο κατά τη διάρκεια του μισού broadcast στο tag team caching

Ευρεστικό Prefetching

- Απλό Ευρεστικό Prefetching
- Εκτελεί έναν υπολογισμό για κάθε αντικείμενο που εκπέμπεται στο κανάλι με βάση την πιθανότητα προσπέλασης P για το αντικείμενο και το ποσό του χρόνου T που θα περάσει μέχρι να εμφανιστεί ξανά η σελίδα στο κανάλι εκπομπής
- Εάν η $P * T$ τιμή των σελίδων που εκπέμπονται είναι υψηλότερες από αυτές των σελίδων στην cache, τότε εκείνες με τη χαμηλότερη τιμή $P * T$, εκδιώκεται από την cache

Ευρεστικό Prefetching

Page	Access Probability	Broadcast Frequency (per Period)	PLX Value
A	P	2	$P/2$
B	$P/2$	2	$P/4$
C	$P/2$	1	$P/2$

Table 2: Access and Frequency Values for the pt Example

Ευρεστικό Prefetching

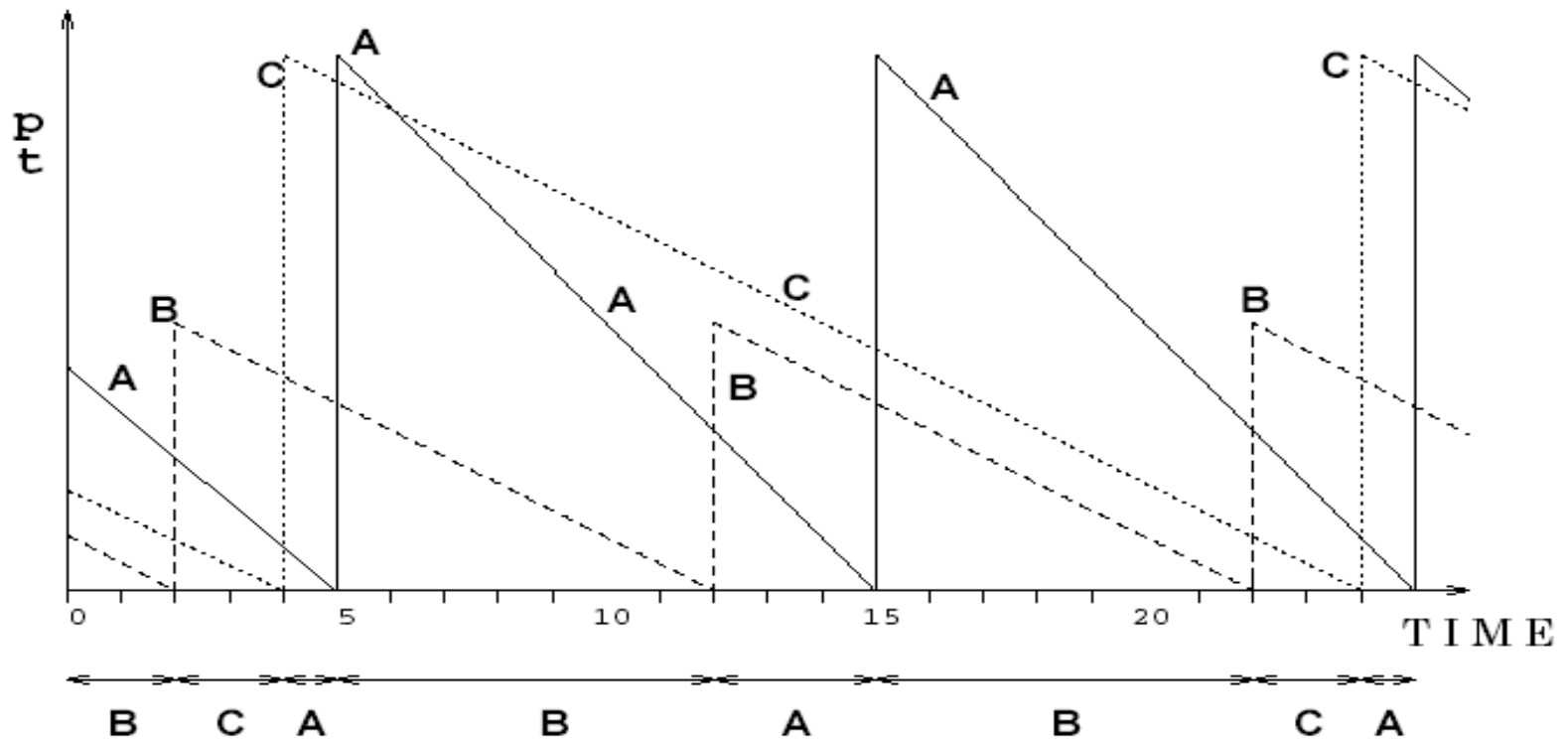


Figure 7: pt Value vs. Time