Κινητός και Διάχυτος Υπολογισμός (Mobile & Pervasive Computing)

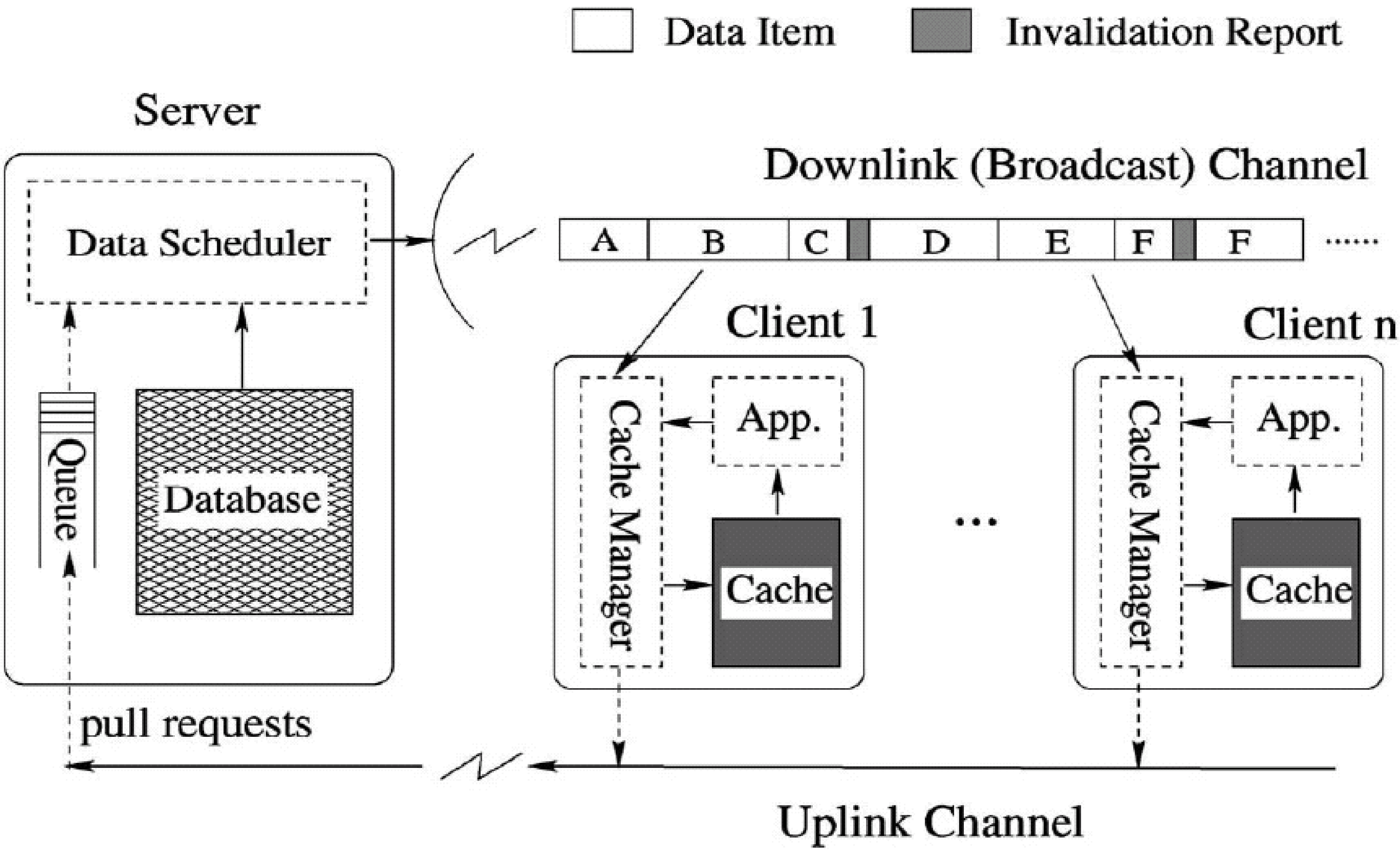
Δημήτριος Κατσαρός

Διάλεξη 10η

Περιεχόμενα

- Πολιτική αντικατάστασης με συνέπεια cache

Αρχιτεκτονική συστήματος



Ορολογία (1/2)

- D: the number of data items in the database.
- C: the size of the client cache.
- $\overline{a_i}$: mean access arrival rate of data item i,

$$i = 1, 2, \ldots, D.$$

- $\overline{u_i}$: mean update arrival rate of data item i,

$$i = 1, 2, \ldots, D.$$

- x_i: the ratio of update rate to access rate for data item i, i.e., $x_i = \overline{u_i}/\overline{a_i}$, $i = 1, 2, \ldots, D$.
- p_i: access probability of data item i, $p_i = \overline{a_i}/\sum_{k=1}^{D} \overline{a_k}$ for $i = 1, 2, \ldots, D$.

Ορολογία (2/2)

- l_i: access latency of data item i, $i = 1, 2, \ldots, D$.
- b_i: retrieval delay from the server (i.e., cache miss penalty) for data item i, $i = 1, 2, \ldots, D$.
- s_i: size of data item i, $i = 1, 2, \ldots, D$.
- v: cache validation delay, i.e., access latency of an effective invalidation report.
- d_k: the data item requested in the kth access,[1] $d_k \in \{1, 2, \ldots, D\}$.
- C_k: the set of cached data items after the kth access, $C_k \subseteq \{1, 2, \ldots, D\}$.
- U_k: the set of cached data items that are updated between the kth access and the $(k+1)$th access, $U_k \subseteq C_k$.
- V_k: the set of victims chosen to be replaced in the kth access, $V_k \subseteq (C_{k-1} - U_{k-1})$.

Η πολιτική Min-SAUD

- Ορίζουμε ένα μέτρο αξίας κάθε αντικειμένου ως εξής:

$$gain(i) = \frac{p_i}{s_i}\left(\frac{b_i}{1+x_i} - v\right)$$

- Μετά από k προσπελάσεις, επιθυμούμε να μεγιστοποιήσουμε τη συνολική αξία των αντικειμένων στην cache, δηλ, να αναγνωρίσουμε το βέλτιστο victim set V_k^*: $V_k^* \subseteq C_{k-1} - U_{k-1}$,

$$V_k^* = arg \min_{V_k \subseteq (C_{k-1} - U_{k-1})} \sum_{i \in V_k} gain(i)$$

$$s.t. \quad \sum_{i \in V_k} s_i \geq \sum_{j \in (C_{k-1} - U_{k-1})} s_j + s_{d_k} - C$$

Optimality της Min-SAUD (1/8)

- Θα δείξουμε ότι η Min-SAUD είναι βέλτιστη για το stretch μέτρο
- Υποθέτουμε Independence Reference Model
- Arrival και Updates είναι Poisson κατανεμημένες
- Interarrival times για αιτήσεις και τροποποιήσεις είναι Εκθετικά κατανεμημένες
- Δηλ., οι αντίστοιχες density functions είναι:

$$f(t_i^a) = \overline{a_i} e^{-\overline{a_i} t_i^a}$$

$$g(t_i^u) = \overline{u_i} e^{-\overline{u_i} t_i^u}$$

Optimality της Min-SAUD (2/8)

- Το access cost είναι το γινόμενο πιθανότητας προσπέλασης επί stretch (access latency προς service time)
- Αγνοώντας το εύρος ζώνης, χρησιμοποούμε το relative access cost, ορισμένο ως πιθανότητα προσπέλασης επί access latency προς μέγεθος, δηλαδή μετά από k προσπελάσεις, έχουμε:

$$S_k = \sum_{1 \leq i \leq D} p_i \cdot \frac{l_i}{s_i}$$

Optimality της Min-SAUD (3/8)

- Θυμηθείτε ότι ακόμα και εάν έχουμε cache hit, δεν μπορούμε να το δώσουμε στην application, παρά μόνο μετά που θα δούμε την IR που περιέχει πληροφορία για το συγκεκριμένο αντικείμενο.
- Έστω ότι $Pr(U_i)$ είναι η πιθανότητα ότι το αντικείμενο i έχει τροποποιηθεί στο διάστημα από την τρέχουσα στιγμή μέχρι την άφιξη μιας τέτοιας IR μετά την query για το i. Η προηγούμενη εξίσωση μπορεί να γραφεί:

$$S_k = \sum_{i \in C_k} \frac{p_i \cdot l_i}{s_i} + \sum_{i \notin C_k} \frac{p_i \cdot l_i}{s_i}$$

$$= \sum_{i \in C_k} \frac{p_i \cdot l_i}{s_i} Pr(U_i) + \sum_{i \in C_k} \frac{p_i \cdot l_i}{s_i} (1 - Pr(U_i)) + \sum_{i \notin C_k} \frac{p_i \cdot l_i}{s_i}$$

Optimality της Min-SAUD (4/8)

- Η εξίσωση της προηγούμενης διαφάνειας αντιστοιχεί σε τρεις περιπτώσεις, τις ακόλουθες:
  1. a cache hit but an obsolete copy,
  2. a cache hit and an up-to-date copy,
  3. a cache miss.
- Η access latency l_i είναι αντίστοιχα:

$$l_i = \begin{cases} v + b_i & \text{if } i \in C_k \text{ and an obsolete copy;} \\ v & \text{if } i \in C_k \text{ and an up-to-date copy;} \\ b_i & \text{if } i \notin C_k. \end{cases}$$

Optimality της Min-SAUD (5/8)

- Θα παράξουμε το Pr(U_i)
- Έστω ότι Pr(U_i') είναι η πιθανότητα τροποποίησης του i από την τρέχουσα στιγμή μέχρι τη στιγμή της επόμενης query για το i.
- Προσεγγίζοντας την Pr(U_i) με την Pr(U_i'), έχουμε:

$$Pr(U_i) \doteq Pr(U_i') = Pr(t_i^u < t_i^a) = \int_{t_i^a=0}^{\infty} \int_{t_i^u=0}^{t_i^a} f(t_i^a) g(t_i^u) dt_i^u dt_i^a$$

$$= \frac{\overline{u_i}}{\overline{u_i} + \overline{a_i}}.$$

- Συνδυάζοντας τις εξισώσεις για το S_k, l_i kai Pr(U_i), έχουμε:

Optimality της Min-SAUD (6/8)

$$S_k = \sum_{i \in C_k} \left(\frac{p_i(v + b_i)}{s_i} \cdot \frac{\overline{u_i}}{\overline{u_i} + \overline{a_i}} \right) + \sum_{i \in C_k} \left(\frac{p_i \cdot v}{s_i} \cdot \frac{\overline{a_i}}{\overline{u_i} + \overline{a_i}} \right) +$$

$$\sum_{i \notin C_k} \frac{p_i \cdot b_i}{s_i} = \sum_{i \in C_k} \frac{p_i (v + \frac{\overline{u_i} b_i}{\overline{u_i} + \overline{a_i}})}{s_i} + \sum_{i \notin C_k} \frac{p_i \cdot b_i}{s_i}.$$

- Με βάση την παραπάνω εξίσωση μπορούμε να δείξουμε ότι:

Theorem 1. *The replacement policy* Min-SAUD *gives better access cost, in terms of stretch, than any other replacement policy.*

Optimality της Min-SAUD (7/8)

- Θα αποδείξουμε την optimality της Min-SAUD, εάν δείξουμε ότι το κόστος S_k είναι πάντα το μικρότερο, εάν χρησιμοποιούμε την πολιτική Min-SAUD. Με επαγωγή
- Έστω ότι S_w είναι το βέλτιστο κόστος για κάποιο k=w για μια άλλη πολιτική
- Έστω ότι V_{w+1} είναι το victim set για να κάνουμε χώρο για το d_{w+1}. Επομένως, έχουμε ότι

$$C_{w+1}=C_w-U_w\cup\{d_{w+1}\}-V_{w+1}$$

- Συνεπώς:

Optimality της Min-SAUD (8/9)

$$S_{w+1} = \sum_{i \in C_{w+1}} \frac{p_i \left(v + \frac{\overline{u_i} b_i}{\overline{u_i} + \overline{a_i}} \right)}{s_i} + \sum_{i \notin C_{w+1}} \frac{p_i \cdot b_i}{s_i}$$

$$= S_w + \sum_{i \in U_w} \left(\frac{p_i}{s_i} \left(\frac{b_i}{1 + x_i} - v \right) \right) - \frac{p_{d_{w+1}}}{s_{d_{w+1}}} \left(\frac{b_{d_{w+1}}}{1 + x_{d_{w+1}}} - v \right)$$

$$+ \sum_{i \in V_{w+1}} \left(\frac{p_i}{s_i} \left(\frac{b_i}{1 + x_i} - v \right) \right)$$

$$= B + \sum_{i \in V_{w+1}} \frac{p_i}{s_i} \left(\frac{b_i}{1 + x_i} - v \right),$$

Optimality της Min-SAUD (9/9)

- Όπου το B είναι ίσο με:

$$B = S_w + \sum_{i \in U_w} \left(\frac{p_i}{s_i} \left(\frac{b_i}{1 + x_i} - v \right) \right) - \frac{p_{d_{w+1}}}{s_{d_{w+1}}} \left(\frac{b_{d_{w+1}}}{1 + x_{d_{w+1}}} - v \right)$$

- Αφού το B δεν ελέγχεται από την πολιτική αντικατάστασης, συμπεραίνουμε ότι το ελάχιστο access cost επιτυγχάνεται όταν διώχνονται τα αντικείμενα με το μικρότερο :

$$\sum_{i \in V_{w+1}} \frac{p_i}{s_i} \left(\frac{b_i}{1 + x_i} - v \right)$$