Data Caching

Cache Consistency Protocol.

**Goal:**
- Guarantee good availability of data for mobile clients.
- Allow some degree of inconsistency
- Fairness among various clients
- Prevent starvation

Diagram:
- **Availability**
  - Stronger
  - Weaker CODA
- **Fairness**
  - AR
  - Other
CODA

Goal: weak consistency, high availability

DS: data server
DC: data client

m: servers
n: clients

DS1 -- DS2 -- DSm

DC1 -- DC2 -- Dcn

either
- strongly connected with the server
- weakly connected w/c.
- disconnected

See pg 12/13

S1 Book

3.4.4 in Disconnected operations

W.C

Heardig S.C

Callback consistency protocol

Client emulate server

S C

W/D.C

Reconnect

Callback

Inform the client of when a data item becomes invalid.
For what kind of applications CoDA protocol is suitable?

Cannot be used for
- chat-like application
- as a lower-level protocol (since it requires user intervention for reintegration)

Mainly suitable for application scenarios where weak consistency can be tolerated and changes of simultaneous updates to same data item are low & availability is of high priority.
Strong consistency protocols.

Broadcasting time-stamped invalidation reports:

Server periodically broadcasts invalidation report which consists of all data streams which were updated in last \( w \) time units (along with when it was updated.

A client waits until the arrival of next invalidation report (IR) to answer the query from an application.

What is the relationship between window size \( w \) and broadcast period \( T \)?
Can $w < T$?

No: Since some update may never be sent to the client, which may lead to inconsistent data access at the client.

$D \geq T$.

What is the how to choose $w$, i.e., how big $w$ should be in comparison to $T$.

Larger the $w$, the longer the period a client (Assume $w = KT$) can be disconnected (or in sleep mode).
duration of sleep
\[ w \geq \frac{ms}{ds} + T \]

\[(w - T) \geq \frac{ms}{ds} \leq \text{max sleep duration}\]

When \( w = T \) the client has to be always connected.

If sleep duration is \( ds \geq \frac{ms}{ds} \) then the local cache needs to be re-validated.

However large \( w \) wasting b/w as well as increases processing time of IT by clients.
T has impact on query response delay. 
T↑ less B/w is consumed however query response delay increases.

The scheme to improve query response delay as well as to reduce B/w wanted for IRs.

Asynchronous → since no periodic broadcast invalidation are broadcast as & when available from the Home agent maintains a "cache state" of the client.

Which items the client is caching.

![Diagram](image.png)
Next Class bring a set of papers which you might be interested in presenting in the class.