Databases

Lesson 03

Data Cache Consistency Maintenance in Mobile and Web Environments

Access Latency in mobile environment

- A device needs a data-record during an application, a request must be sent to the server for the data record– mechanism called pulling
- **Access latency**– the interval for the application software to access a particular record

Data cache maintenance necessity in mobile environment

- Caching and hoarding the record at the device reduces access latency to negligible

Data Cache inconsistency

- Means that data records cached for applications are not invalidated at the device when modified at the server but not modified at the device

Methods for Cache consistency Maintenance

1. Cache invalidation mechanism (server-initiated case): the server sends invalidation reports on invalidation of records (asynchronous) or at regular intervals (synchronous)

Methods for Cache consistency Maintenance

2. Polling mechanism (client-initiated case): Polling means checking from the server, the state of data record whether the record is in valid, invalid, modified, or exclusive state

Polling Method

- Each cached record copy polled (checked) whenever required
- The device connects to the server and finds out data record copy at the device become invalid or has been modified at the server
- If modified or invalidated, then the device requests for the modified data and replaces

Method of Time-to-live mechanism (client-initiated case)

- Each cached record assigned a TTL (time-to-live)
- The TTL assignment is adaptive (adjustable) previous update intervals of that record
- After the end of the TTL, the client device polls for invalid or modified state

Method of Time-to-live mechanism (client-initiated case)

- If invalid or modified state found then the device requests the server to replace
- When TTL is set to 0, the TTL mechanism is equivalent to the polling mechanism

Cache consistency Maintenance by Invalidation Mechanisms

- If the database server changes a data record (either read-only or writable), of which one or more client devices have copies (read-only), then the server-record invalidated, forcing all client devices that currently have copies of the record to give up (not use further in computations) their cache-copies

Four Cache Invalidation Mechanisms

- Stateless asynchronous
- Stateless synchronous
- Stateful asynchronous
- Stateful synchronous

Stateless asynchronous

- Device cache states are not maintained at the server
- The server advertises invalidation report
- On receiving an invalidation report, the client requests for or caches the new data record

Advantage of the asynchronous approach

- No frequent, unnecessary transfers of data reports
- Mechanism more bandwidth efficient

Disadvantages of this approach

- Every client device gets an invalidation report, whether that client requires that copy or not
- Client devices presume that as long as there is no invalidation report, the copy is valid for use in computations

Disadvantages of this approach

- Therefore, even when there is link failure, the devices may be using the invalidated data and the server is unaware of state changes at the clients after it sends the invalidation report

Stateless synchronous

- Device cache states are not maintained at the server
- The server periodically advertises invalidation report
- The client requests for or caches the data on receiving the invalidation report
- If no report is received till the end of the period, the client requests for the report

Advantage of the synchronous approach

- Client devices receive periodic information regarding invalidity (and thus validity) of the data caches
- The periodic invalidation reports lead to greater reliability of cached data as update requests for invalid data can be sent to the server by the device-client. This also helps the server and devices maintain cache consistency through periodical exchanges

Disadvantage of the synchronous approach

- Unnecessary transfers of data invalidation reports take place
- Every client device gets an advertised invalidation report periodically, irrespective of whether that client has a copy of the invalidated data or not

Disadvantage of the synchronous approach

- During the period between two invalidation reports, the client devices assume that, as long as there is no invalidation report, the copy is valid for use in computations

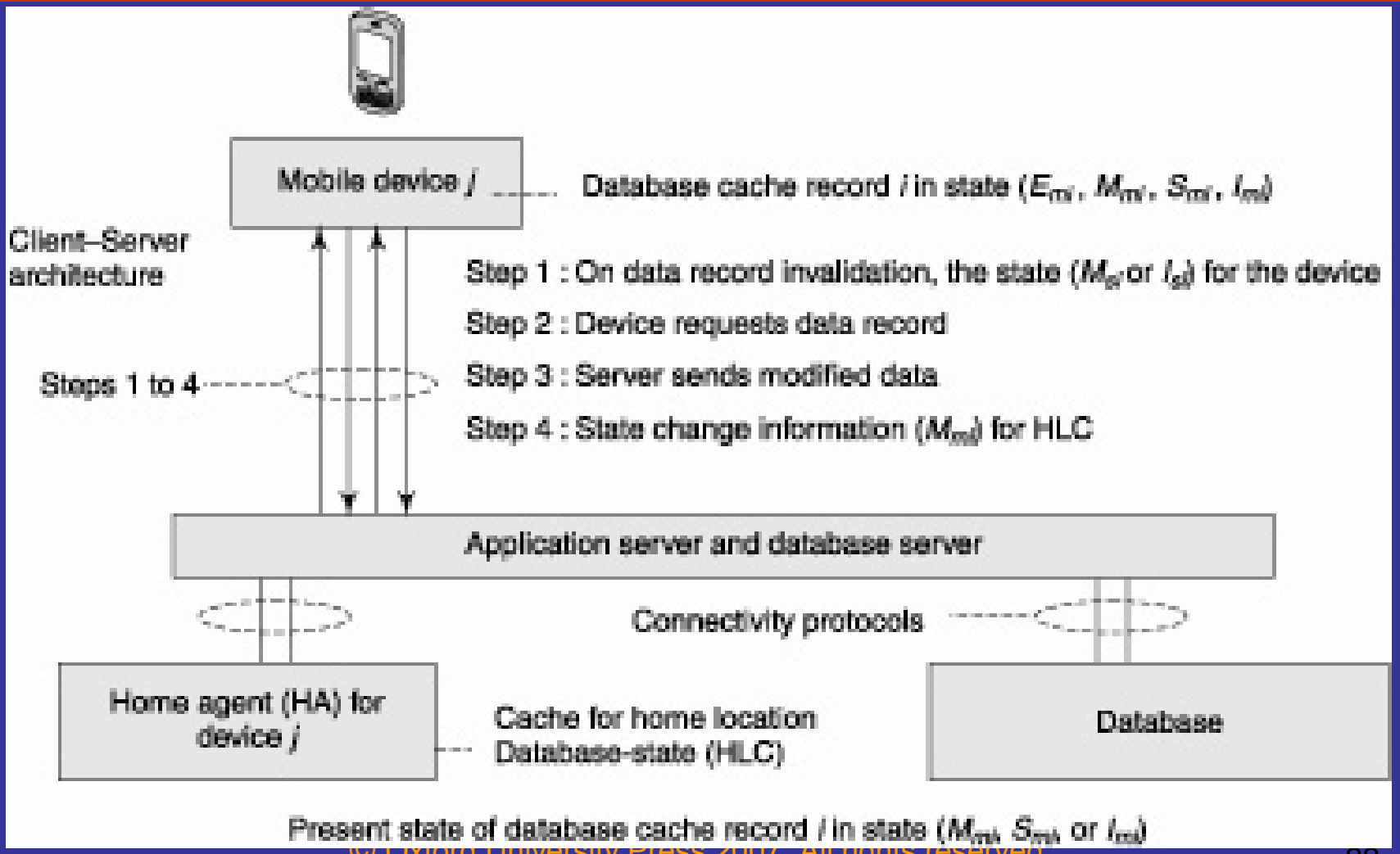
Stateful asynchronous

- Device cache states are maintained at the server
- The server transmits invalidation report to concerned devices
- The client requests for new data on receiving the invalidation report and sends its new state after caching new record from the server

Advantage of the stateful asynchronous approach

- Server keeps track of the state of cached data at the client device
- This enables the server to synchronize with the state of records at the device cache and keep the home locating cache (HLC) updated

Four-step stateful asynchronous cache invalidation mechanism for cache consistency



Advantage of the stateful asynchronous approach

- The stateful asynchronous mode is also advantageous in that only the affected clients receive the invalidation reports and other devices are not flooded with irrelevant reports

Disadvantage of the Asynchronous stateful

- Client devices presume that, as long as there is no invalidation report, the copy is valid for use in computations.
- Therefore, when there is a link failure, then the devices use invalidated data

Stateful synchronous

- Device cache states maintained at the server
- The server periodically transmits invalidation report to the concerned devices

Stateful synchronous

- The client requests for new data on receiving the invalidation report, and sends its new state after caching the new data.
- If no report is received till the end of the period, the client requests for the report.

Advantage of the stateful synchronous approach

- Reports identifying invalidity (and thus, indirectly, of validity) at periodic intervals and that the server also periodically updates the client-cache states stored in the HLC
- Enables to synchronize with the client device when invalid data gets modified and becomes valid.

Advantage of the stateful synchronous approach

- Each client be periodically updated of any modifications at the server.
- When the invalidation report is not received after the designated period and a link failure is found at the device, the device does not use the invalidated data
- Instead it requests the server for an invalidation update.

Disadvantage of the stateful synchronous approach

- High bandwidth requirement to enable periodic transmission of invalidation reports to each device and updating requests from each client device

Web cache maintenance in mobile environments

- Mobile devices or their servers can be connected to a web server (for example, traffic information server or train information server)
- Web cache at the device stores the web server data and maintains it in a manner similar to the cache maintenance for server data

Web cache maintenance in mobile environments

- Mobile devices or their servers connected to a web server— for example, traffic information server or train information server
- Web cache at the device stores the web server data and maintains it in a manner similar to the cache maintenance for server data

Web cache maintenance necessity

- In a mobile environment to overcome access latency in downloading from websites due to disconnections

Two methods for Web cache consistency maintenance

- Time-to-live (TTL) mechanism (client-initiated case)
- Power-aware computing mechanism (client-initiated case): Each web cache maintained at the device can also store the CRC (cyclic redundancy check) bits

Power Aware Method

- Assume that there are N cached bits and n CRC bits and N >> n
- Server also stores *n* CRC bits
- Consistency between the server and device records, the CRC bits at both identical
- Else the corresponding CRC bits at the server modifies.

Power Aware Method

- After the TTL expires or on-demand for the web cache records by the client API, the cached record CRC polled and obtained from the website server

Power Aware Method

- If the *n* at server found to be modified and the change is found to be much higher than a given threshold (i.e., a significant change), then the modified part of the website hypertext or database is retrieved by the client device for use by the API

Summary

- Four methods of Cache Invalidation based cache consistency maintenance mechanism
  - Stateless asynchronous
  - stateless synchronous
  - stateful asynchronous
  - stateful synchronous mechanisms

...

Summary

- Polling and TTL methods for cache consistency maintenance for data cache
- Similar methods for Web Cache
- Power aware method– Poll for significant change in CRC

End of Lesson 03

Data Cache Consistency Maintenance in Mobile and Web Environments