Department of Computing

**Hong Kong Polytechnic University**

**Comp 5527 Mobile Computing and Data Management**

**Tutorial Four**

1. Broadcast disk caching is different from traditional caching. It considers both a data item’s access probability and broadcast frequency. PIX cost function is one metric to determine data replacement in broadcast disk caching. It is defined as:

cost = access probability / broadcast frequency

Given the following data set

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Group | A | B | | C | | | | | | | |
| Data | 2 | 5 | 8 | 1 | 12 | 15 | 18 | 22 | 25 | 28 | 31 |
| 9 | 6 | 10 | 3 | 13 | 16 | 19 | 23 | 21 | 29 | 32 |
| 26 | 7 | 11 | 4 | 14 | 17 | 20 | 24 | 27 | 30 | 33 |
| Access probability | 4/21 | 1/21 | 1/21 | 1/168 | 1/168 | 1/168 | 1/168 | 1/168 | 1/168 | 1/168 | 1/168 |

we designed a broadcast cast schedule as following

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 2 | 5 | 1 | 12 | 2 | 8 | 15 | 18 | 2 | 5 | 22 | 25 | 2 | 8 | 28 | 31 |
| 9 | 6 | 3 | 13 | 9 | 10 | 16 | 19 | 9 | 6 | 23 | 21 | 9 | 10 | 29 | 32 |
| 26 | 7 | 4 | 14 | 26 | 11 | 17 | 20 | 26 | 7 | 24 | 27 | 26 | 11 | 30 | 33 |

Assuming that the cache is of size 3, and a data item access sequence is 2, 5, 4, 7, 5, 3, 5, 9, please describe the change of cache status.

1. Mobile wireless networks can be broadly classified as either infrastructure based or ad hoc based as shown in the figure 1 and figure 2 respectively. Cache consistency has been considered in many works of infrastructure based mobile network. Explain why the approaches used in infrastructure based mobile network cannot be directly used in ad hoc based mobile network.



Fig 1 Infrastructure-based Mobile Network Fig 2 Ad Hoc based Mobile Network

1. In cache placement, a node needs to determine to cache data or cache the path accessing the data.

In “cache data” schema, 1) the data item *di* is at least cached at the requesting node; 2) a node caches the passing-by data *di* locally when it finds *di* is popular, i.e., there were many requests for it; 3) a node does not cache the passing-by data *di* if it finds all requests for *di* are from the same node.

In “cache path” schema, 1) the data item *di* is at least cached at the requesting node; 2) a node only records the data path when it is closer to the caching node than the data center.

The following is an ad hoc mobile network and node 11 is the data source. Given the different data accessing scenarios, please indicate which nodes need to cache the data/path.

a) In data cache schema: both N6 and N7 request a data di through N5,

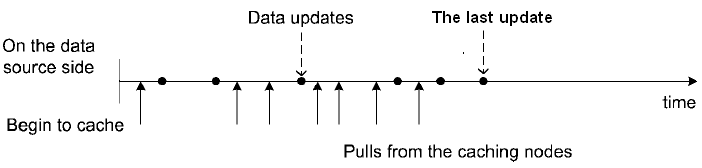
b) In data cache schema: both N1 and N2 request a data di through N3-N4-N5

c) In data cache schema: N1 request a data di through N3-N4-N5

d) In path cache schema: N1 request a data di through N3-N4-N5



1. (After class exercise) In Predictive Caching Consistency protocol (PCC), the source node does not push the update to the caching nodes every time a data item is updated. Instead, it pushes only when at least *Th* caching nodes will pull this data item. The source node predicts the expected number of pulls between the latest and the forthcoming update based on recorded history. Given the following history, for *Th*=*1.5* and *2* respectively, explain whether push is needed at the last update.



We use *p(v|h)* to denote the probability that given the history *h*, a number of *v* caching nodes will pull the data. You may use the following information: p(0|102310)=0.3 p(1|102310)=0.3, p(2|102310)=0.1, p(3|102310) =0.1; p(0|112311)=0.4, p(1|102311)=0.4, p(2|102311)=0.05, p(3|102311) =0.05.