Advanced Systems Topics 2004 – Paper 9 Answer 5 (JAC)

This material was covered in JAC's last lecture in AST in Lent term, as well as being easily available in handouts and in the book suggested for reading, and partly inherent in the guest lecture from MSR.

(*a*) Describe the DHT algorithms that are used in the CAN and Pastry peer-to-peer system designs to route a request for content to the node that holds that content. In your description, you must also cover how nodes are added and deleted. [5 marks each]

Salient points in answer must include: Background - DHT puts content on a logical node by hashing key onto logical node i-d. . Need to organise nodes to navigate too.

CAN - uses multiple hashes on content to create n-D index, then traverses the n-D space by taking "diagonal" Path through nodes. Pastry uses prefix routing.

Node addition in CAN involves random placement of node in $n - D$ space, and then partition of the hypercube that is held by the nearest node in that space and assignment of appropriate half content to the new node. Navigation is simply by neighbour nodes along "diagonal" in the CAN $n - D$ space.

Node addition in Pastry involves organising nodes in a Ring (like Chord) and then uses prefix routing, finding the next node around the "circle" of 2^n nodes by going to the nearest one that is there in lg base m bits size steps. nodes keep sets of pointers to forward and backward nodes in this ring. Addition and deletion is simply like adding nodes in a doubly linked list.

(*5*) CAN and Pastry have proposed mechanisms to reduce the latency of search. Describe the ways these work, and discuss their relative complexity and performance. [5 marks each]

Both use pings (ICMP echos) to find the distance from candidate neighbours. CAN uses a technique called binning, where nodes are clustered in geographically close regions, and the node id space warped to match. Pastry does similar, but on the ring id.

The base line complexity of the schems are logarithmic in the patrh length to get from hash(key) to node, and without proximity, this means that they are order this in latency to get to a node.

Both depend partly on accuracy of ping as estimate. The big problems are that this may not match real proximity, and also that the Internet is not a euclidean plane.