

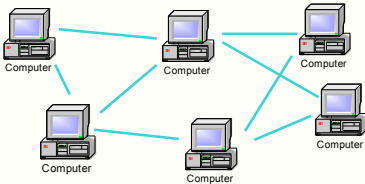
Mei Li Wang-Chien Lee Anand Sivasubramaniam

Introduction

Problem: massive amount of information in Internet

- ◆ naturally distributed in wide area
- ◆ centralized collection undesirable
- ◆ dynamic and user-specific queries

Internet-scale semantic-based search



Objective: Design a peer-to-peer overlay supporting semantic-based search in an environment that has:

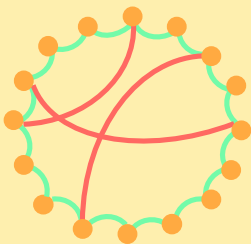
- ◆ large scale
- ◆ dynamic changes
- ◆ no central control
- ◆ high dimensionality

Proposal

Semantic Small World

- ◆ Position peers and data objects in the semantic space.
- ◆ Peers with similar data objects form into peer clusters.
- ◆ Data space is linearized.
- ◆ Peer clusters form into small world over the linearized space.

Background: Small World Network



- Regular short range contacts ensure connectivity
- Random long range contacts provide shortcut

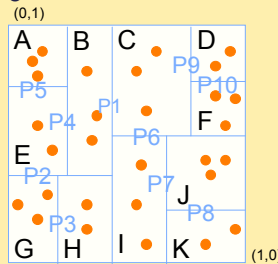
Semantic based peer positioning

A peer clusters its local data and chooses the centroid of its largest data groups as the position to join the network.

Cluster based data management

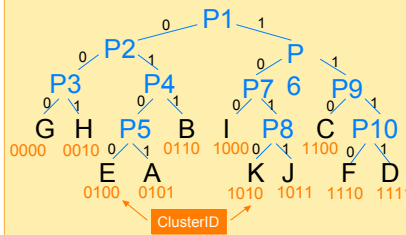
Peers form into clusters and share the responsibility of taking charge of data subspaces.

Data space is adaptively partitioned according to data distribution.



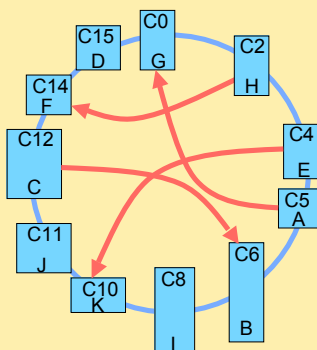
Adaptive space linearization

Peer clusters and data subspaces are linearized during peer join and space partition.

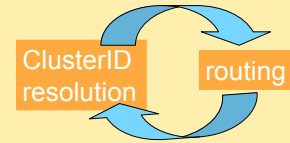


Small world network construction

One dimensional small world network is constructed over the linearized ClusterID space.



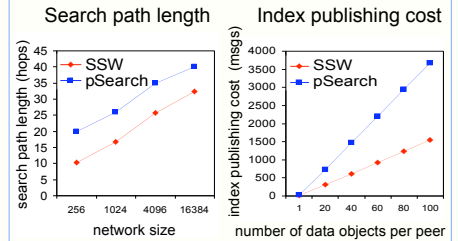
Search in SSW



Average path length: $O(\log^2 N/L)$

Performance

- ◆ efficient search
- ◆ low maintenance cost
- ◆ adaptive to data distribution and query distribution
- ◆ fair load balance
- ◆ resilient to failure



Summary

SSW mimics the information flowing model in social network to support efficient information flow in networks arising in technology.

Applications

Platform for Internet-scale information centric applications

- resource discovery
- file sharing
- network monitoring
- content-based multicast

Conclusion

- ◆ We designed SSW, an efficient and robust overlay for P2P systems.
- ◆ SSW can be applied for a variety of applications, ranging from file sharing to content-based multicast in large scale dynamic networks.