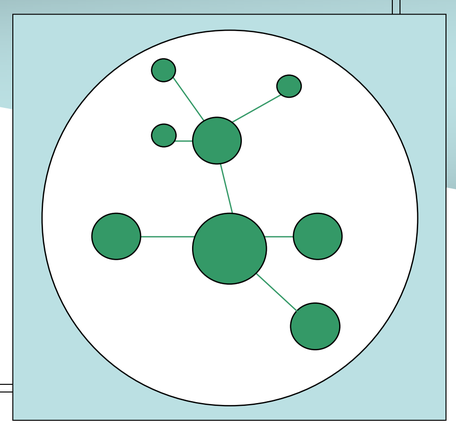
HyperCuP – A Scalable and Ontology-Based Peer-to-Peer Infrastructure for Semantic Web Services

Mario Schlosser[1], Michael Sintek[1,2], Stefan Decker[1], Wolfgang Nejdl[1,3]

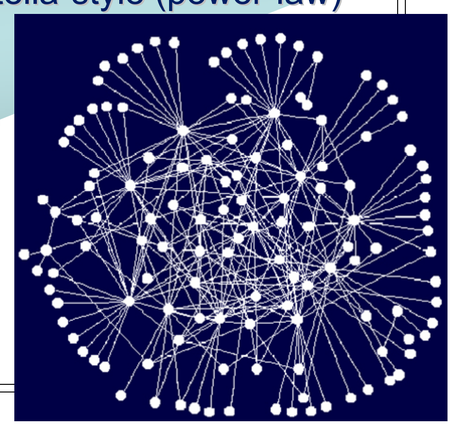
Semantic Service Web

- Large network of service providers capable of instantiating high-level task descriptions in a distributed fashion
- Task description to be broadcasted among possibly interesting service providers
- Services classified by global service ontology

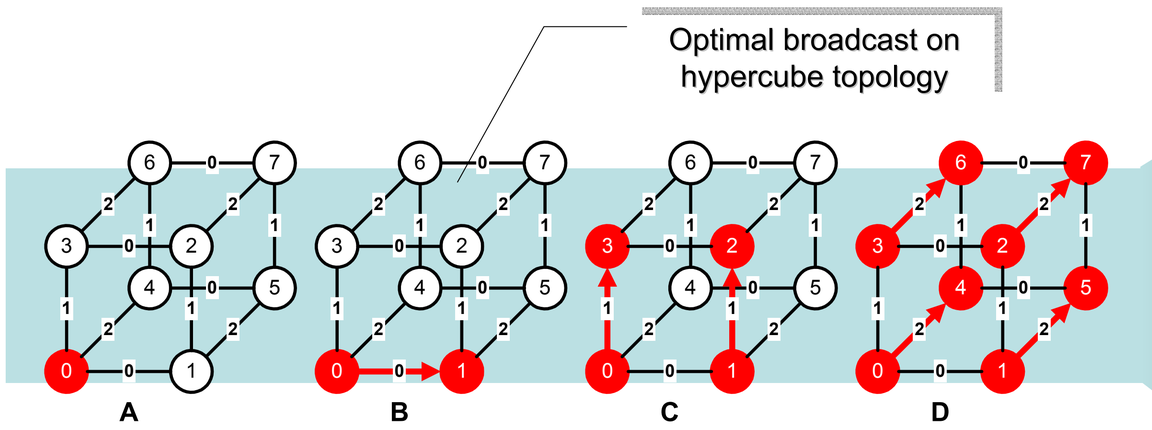


Peer-to-Peer Networks

- Decentralized, real-time, immediate
  - Best fit for dynamic service networks
- Evolution of current Gnutella-style (power-law) networks
  - Poor scalability to large number of nodes
  - Incomplete search without guarantees
  - Long search times
  - Vulnerable against malicious attacks



Deterministic and Semantically Organized Network Topology



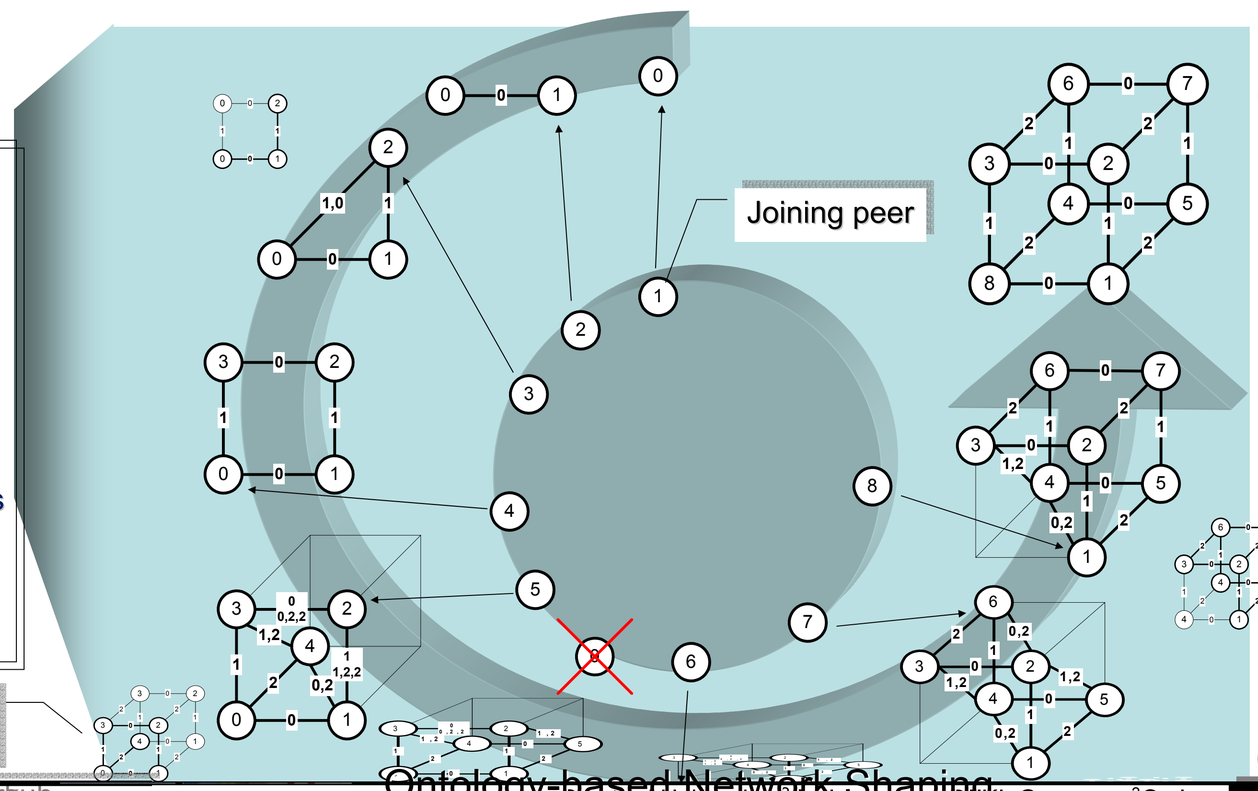
Hypercubes: A Deterministic P2P Topology

- Interesting topology properties
  - Logarithmic diameter – bounded search times
  - Vertex symmetry – load balancing over nodes
  - Fault tolerance – protection against attacks
- Efficient search and broadcast algorithms
  - Optimal broadcast, one message per peer
  - Shortest path routing to any peer

1 Network Topology

Distributed Topology Construction

- Properties
  - Decentralized: No central servers, no super peers
  - Message and time complexity: Logarithmic to number of peers in the network
  - Self-healing: Recovery from node failures
- Algorithm
  - Always implicitly maintain topology of next biggest hypercube in node connections
  - Remaining nodes cover positions of departing nodes based on deterministic buffering scheme
  - Arriving nodes are able to join via any peer in the network and take over position(s) in the hypercube



2 Ontology-Based Network Construction

Ontology-Based Routing

Concept-driven peer clustering

- Hypercube network is decomposed into concept clusters, containing peers associated with a particular combination of ontology concepts supported by those peers
- Concept clusters are sub-hypercubes of top hypercube
- Addressing scheme: Ontology concept coordinates represent logical conjunction of supported concepts, storage coordinates allow for multiple peers per cluster

Algorithms

- Shortest path routing on ontology coordinates
- Flooding optimal broadcast on storage coordinates to reach all peers in concept sub-hypercubes
- Restricted optimal broadcast on ontology coordinates and logic minimization to answer queries consisting of conjunctions and disjunctions of ontology concepts

- Scales to millions of peers while allowing for complex queries and providing search guarantees at optimal routing behavior

[1]DFKI, Germany [2]Technical University Hannover, Germany [3]On leave from DFKI, Germany

P2P.semanticweb.org